

**FINAL TECHNICAL REPORT: USGS AWARD
G20AP00054**

**A Partially Non-Ergodic Ground-Motion Model for Cascadia Interface
Earthquakes**

Chih-Hsuan Sung

University of California, Berkeley

Norman Abrahamson

SC Solutions

Grant Period: May 2020 - April 2021

Submitted on NOV 2, 2021

This material is based on work supported by the U.S. Geological Survey (USGS), Department of the Interior, under USGS award number (G20AP00054). The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the opinions or policies of the U.S. Geological Survey. Mention of trade names or commercial products does not constitute their endorsement by the U.S. Geological Survey.

Abstract

New ground-motion models (GMMs) for subduction zones were developed as part of the NGA-Sub project. Three of these models included regional differences in the ground motions for seven regions, including the Cascadia region. Within the Cascadia region, the models are ergodic; however, non-ergodic effects on the ground motions in Cascadia can be seen in the suite of numerical simulations of ground motions from megathrust earthquakes on the Cascadia subduction zone through a 3-D crustal structure developed by the M9 project (Frankel et al., 2018). The Abrahamson and Gulercce (2020) GMM (AG20) is modified to include the non-ergodic effects from the 3-D simulations. First, the scaling of the basin effects as a function of the depth to a shear-wave velocity of 2.5 km/s ($Z_{2.5}$) in the AG20 model is modified to be consistent with the $Z_{2.5}$ scaling from the 3-D simulations. Second, the spatial distribution of the non-ergodic site terms is estimated using the varying coefficient model for the region covered by the 3-D velocity model. The integrated nested Laplace approximation is used to handle the very large data set from the simulations. The site terms are estimated for the long-period range (1-10 sec) that reflects the period range of the deterministic part of the M9 simulations. With the non-ergodic site terms, the aleatory variability (single-station sigma) for the 3-D simulations reduced by 15-25% compared to an ergodic standard deviation for Cascadia. In addition to an average single-station sigma, a spatially varying single-station sigma model is developed which shows the highest variability for sites near the basin edges. The epistemic uncertainty in the non-ergodic site terms for a single 3-D velocity model is small, but there will be uncertainty due to alternative 3-D models. Without simulation results for different 3-D velocity models, we set the epistemic uncertainty to be one half of the between-site standard deviation from the simulations. The resulting non-ergodic GMM can be used in seismic hazard analyses that include site-specific basin effects for the median and aleatory variability for interface events in the Seattle region. As an example application, the seismic hazard for interface events using the non-ergodic GMM is compared to the hazard using the original AG20 ergodic GMM. For the example sites, the non-ergodic GMM leads to

significant changes in the $T=3$ sec hazard for specific sites. The changes in the $T=3$ sec hazard from interface events only lead to expected increases and decreases of 2500-yr ground motion for the total hazard (from all sources) of up to a factor of 1.25.

1 INTRODUCTION

2 Over the last decade, there has been a move from using global average ground-motion models
3 (GMMs) in probabilistic seismic hazard (PSHA) to using non-ergodic GMMs that account
4 for the systematic and repeatable effects for broad regions. The recently developed Next Gen-
5 eration Attenuation - Subduction (NGA-SUB) GMMs by Abrahamson and Gulerce (2020),
6 Kuehn et al. (2020), and Parker et al. (2020) are examples of regionalized GMMs for sub-
7 duction zone earthquakes. These three GMMs include regional differences for seven broad
8 regions: Japan, Taiwan, Cascadia, Alaska, Mexico, Central America, South America, and
9 New Zealand. The GMMs include regional differences in the site term, linear distance term,
10 basin term, constant term, and break point for large-magnitude scaling. With these regional
11 differences, the NGA-SUB GMMs capture the average differences in the ground-motion scal-
12 ing between regions.

13 Within each broad region, the NGA-SUB GMMs use the ergodic assumption; however,
14 there can be significant differences within a region. The development of non-ergodic effects
15 on GMMs for small regions or for specific sites and earthquake locations requires source/site-
16 specific data. For the Cascadia region, empirical ground-motion data from interface earth-
17 quakes are not available to constrain the non-ergodic terms. With the lack of empirical
18 ground-motion data from Cascadia interface earthquakes, 3-D simulations can be used to
19 constrain the repeatable path and site effects within the region.

20 The 3-D simulations from the M9 project (Frankel et al., 2018) showed that there are
21 large systematic differences in the site and path effects due to the 3-D velocity structure in
22 the Cascadia region. The results from the M9 project have been used to develop long-period
23 basin factors for ground motions from interface events in the Seattle region using a reference-

station approach. For periods greater than 2 sec, the Seattle Department of Construction and Inspection (SDCI) specifies a basin factor of 2 relative to a reference site with $Z_{2.5}=3000$ m (SDCI, 2018).

In this study, we use the 3-D simulations from the M9 project to modify the NGA-SUB GMM developed by Abrahamson and Gulerce (2020) (called AG20). First, the basin scaling in the AG20 model is replaced with basin scaling that is based on the M9 simulations. Second, the AG20 model with the new basin scaling is modified to be a partially non-ergodic GMM based on the combined path and site effects determined from the M9 project 3-D simulations. Models for reduced single-station sigma are developed including an average value that applies to all sites and site-specific values that captures the spatial variability of the standard deviation. Example hazard calculations for the interface source are shown using the AG20 ergodic GMM and the modified AG20 non-ergodic GMM.

Most of the features of basin scaling shown in this study have been previously identified by Frankel et al. (2018) in their evaluation of the M9 simulations. The main difference in this study is that we incorporate these features into a non-ergodic version of the AG20 GMM in a manner that is consistent with the development of GMMs for use in seismic hazard calculations. We also address all the issues for the application of the proposed non-ergodic GMM in probabilistic seismic hazard calculations.

AG20 CASCADIA GROUND-MOTION MODEL

The AG20 GMM was developed based on the subset of the NGA-SUB database that included 4118 recordings from 113 subduction interface earthquakes ($M_w = 5.0 - 9.2$) and 4850 recordings from 89 intra-slab events ($M_w = 5.0 - 7.8$) in the global data set. For Cascadia, the AG20 data set includes only 177 recordings from 6 intra-slab events. For interface earthquakes, the AG20 Cascadia model uses the average world-wide scaling between interface and intra-slab earthquakes to estimate the interface ground motions from the available intra-slab

ground motions in Cascadia. In the AG20 model, the difference between ground motions for interface and intra-slab earthquakes is not just a constant factor. The magnitude, distance, and depth scaling are also different for interface and intra-slab earthquakes, so the scale factor between interface and intra-slab earthquakes depends on magnitude, distance, and depth. Examples of the scale factors between interface and intra-slab for $T=0.2$ and $T=2$ sec from the AG20 GMM, based on global data, are shown in Figure 1. The factors range from 1.3 at 50 km distance to 2.5 at 500 km distance. These factors were used in the AG20 GMM to estimate the interface ground motions for Cascadia from the available intra-slab data in the Cascadia region.

The AG20 model, without basin-depth scaling, has the following functional form:

$$\begin{aligned} \ln(PSA) = & a_1 + (a_2 + a_3(M - 7)) \ln(R_{RUP} + f_{FF}(M)) \\ & + f_{mag}(M, F) + f_{Z_{TOR}}(Z_{TOR}) + f_{site}(PGA_{1000}, V_{S30}) \\ & + f_{slab}(R_{RUP}) + a_{15}F_{AS} \\ & + RS_{region}(V_{S30}, R_{RUP}, Z_{2.5}) + ADJ_{CAS} \end{aligned} \quad (1)$$

in which $PSA(g)$ is the $RotD_{50}$ 5%-damped pseudo spectral acceleration (PSA) in g; M is the moment magnitude; R_{RUP} is the shortest distance from the site to the rupture plane in km; $f_{FF}(M)$ is the finite-fault term; F is the event type (0 for interface and 1 for intra-slab); Z_{TOR} is the depth to the top of the rupture plane in km; V_{S30} is the time-averaged shear-wave velocity over the top 30 meters in m/sec; PGA_{1000} is the median peak acceleration (g) for $V_{S30} = 1000$ m/sec; F_{AS} is a dummy variable (0 for mainshock and 1 for aftershock). The RS_{region} term contains the region-specific terms for the seven regions, including the constant term, linear-R term, V_{S30} term. The AG20 GMM was first developed without the basin term because many of the sites in the NGA-SUB data base does not have estimates of the basin depth.

The AG20 model includes two alternative models for Cascadia: the first model is based on the sparse ground-motion data in the region; the second model is scaled to increase the amplitudes of the short-period ground motion to be more consistent with other regions. This adjustment is given by the ADJ_{CAS} term in equation (1). The "adjusted" Cascadia model is used as the base ergodic model in the current study. In the following, we call the AG20 GMM for Cascadia, including the ADJ_{CAS} term, the AG20 model for simplicity.

The AG20 model includes the Cascadia-specific basin scaling based on the depth to the 2.5 km/sec shear-wave velocity horizon below the site, $Z_{2.5}$. In the AG20 dataset for Cascadia, there is a correlation between the V_{S30} and $Z_{2.5}$ values (Figure 2). If the independent parameters are correlated in empirical data sets, there can be a high correlation between the estimated coefficients. To reduce the trade off between the coefficients for the $Z_{2.5}$ and V_{S30} terms in the regression, the AS20 model uses a normalized basin depth term, $Z'_{2.5}$, given by:

$$Z'_{2.5} = \frac{Z_{2.5} + 50m}{Z_{2.5ref}(V_{S30}) + 50m} \quad (2)$$

in which $Z_{2.5ref}(V_{S30})$ is the region-specific median $Z_{2.5}$ for a given V_{S30} . The 50 m is added so that the log term does not increase too rapidly for small $Z_{2.5}$ values. An important additional advantage of this formulation is that it allows basin-depth scaling from 3-D simulations to be incorporated into the GMM in a manner that is consistent with the V_{S30} scaling in the GMM as discussed later.

In the AG20 model, the $Z_{2.5ref}$ for Cascadia is modeled by a tri-linear relation shown in Figure 2.

$$\ln(Z_{2.5ref}(V_{S30})) = \begin{cases} 8.52 & \text{for } V_{S30} \leq 200 \text{ m/s} \\ 8.52 - 0.88 \ln(V_{S30}/200) & \text{for } 200 \text{ m/s} < V_{S30} < 570 \text{ m/s} \\ 7.6 & \text{for } V_{S30} \geq 570 \text{ m/s} \end{cases} \quad (3)$$

The evaluation of the basin scaling in the AG20 subset of ground-motion data from Cascadia found a strong increase in the long-period ground motion with increasing basin depth for the deep basin sites ($Z_{2.5} > 2000$ m), but the available ground-motion data for shallow basin sites in Cascadia ($Z_{2.5} < 2000$ m) did not indicate a clear dependence on basin depth. Therefore, the AG20 model used a bi-linear model for the functional form of the basin-depth scaling with no basin-depth effects for $Z'_{2.5} < 1$ as shown in equation (4).

$$f_{basin} = \begin{cases} a_{39} \ln(Z'_{2.5}) & \text{for } Z'_{2.5} > 1 \\ 0 & \text{otherwise} \end{cases} \quad (4)$$

With the $Z_{2.5ref}$ for Cascadia, this corresponds to no basin scaling for $Z_{2.5}$ less than about 2000 m for soft-rock sites and for $Z_{2.5}$ less than about 5000 m for soft-soil sites. This does not mean that there is no effect of the basin for soft-soil sites for $Z_{2.5}$ less than 5000 m, but rather that the V_{S30} scaling in the GMM already accommodates this scaling due to the correlation between V_{S30} and $Z_{2.5}$ in the empirical data used to develop the AG20 model.

Constraining Basin Scaling in GMMs Using 3-D Simulations

Simulations using 3-D velocity models often limit the shallow V_S values to reduce computational time and because the available geophysical and seismic data used to develop the 3-D velocity structure do not provide constraints on the shallow V_S profile. In the M9 project, the shallow V_S profile is set at $V_{S30}=600$ m/s for the full region. When using 3-D simulations with the same shallow V_S profile for all sites, then there is no V_{S3-0} dependence, and the site scaling in the simulations can be modeled as a function of the basin depth parameter, $Z_{2.5}$, by itself. For example:

$$C_{basin}(Z_{2.5}) = c_0 + c_2 \ln(Z_{2.5}) \quad (5)$$

In empirical GMMs, the site scaling is often developed for the V_{S30} term first, without

108 the $Z_{2.5}$ scaling. This approach was used in the development of the AG20 GMM: the V_{S30}
 109 scaling was estimated and smoothed over period, then the $Z_{2.5}$ scaling was estimated with
 110 V_{S30} scaling fixed to the smoothed model. Due to correlation between V_{S30} and $Z_{2.5}$ in the
 111 empirical data set, the sites with lower V_{S30} values have larger $Z_{2.5}$ values on average. As a
 112 result of this correlation, the V_{S30} scaling in the AG20 GMM accounts for some of the $Z_{2.5}$
 113 scaling. If we took the scaling from 3-D simulations as given by equation (5) and simply
 114 added this to GMM with the V_{S30} scaling fixed from the regression, then we would double
 115 count the part of the $Z_{2.5}$ scaling that was mapped into the V_{S30} scaling in the GMM. This
 116 double counting can be avoided by using the normalized basin depth term, $Z'_{2.5}$. With this
 117 normalized term, equation (5) can be written as:

$$\begin{aligned}
 C_{basin}(Z_{2.5}, V_{S30}) &= c_0 + c_2 \ln(Z'_{2.5}) \\
 &= c_0 + c_2 (\ln(Z_{2.5})) - c_2 (\ln(Z_{2.5ref}(V_{S30})))
 \end{aligned} \tag{6}$$

118 If there is a correlation between $Z_{2.5}$ and V_{S30} in the empirical data set, then the $Z_{2.5ref}$ will
 119 be a function of V_{S30} . The last term in equation (6) removes the part of the $Z_{2.5}$ scaling that
 120 is already explained by the V_{S30} scaling in the GMM. If there is no correlation between $Z_{2.5}$
 121 and V_{S30} , then the $Z_{2.5ref}$ will be independent of V_{S30} , and the last term is just a constant.
 122 Using the normalized basin depth, the scaling on V_{S30} is adjusted back to the value for the
 123 shallow site effect and the scaling on the $Z_{2.5}$ is for the full basin effect.

124 To show how this works, consider that the shallow site amplification and the deep basin
 125 amplification are independent physical effects and that the true site term is given by:

$$f_{site}(Z_{2.5}, V_{S30}) = c_0 + c_1 \ln(V_{S30}) + c_2 \ln(Z_{2.5}) \tag{7}$$

126 Next, assume that in the empirical data set, the V_{S30} and $Z_{2.5}$ are correlated, such that the
 127 $Z_{2.5ref}$ is given by:

$$\ln(Z_{2.5ref}) = a_1 + a_2 \ln(V_{S30}) \tag{8}$$

128 If the GMM is first developed using a regression with V_{S30} scaling, without the $Z_{2.5}$ scaling,
 129 then the site term is given by:

$$f_{site}(Z_{2.5}, V_{S30}) = c'_0 + c'_1 \ln(V_{S30}) \quad (9)$$

130 in which $c'_1 = c + 1 + c_2 a_2$. The slope of $\ln(V_{S30})$ is different from the true c_1 value because
 131 some of the $\ln(Z_{2.5})$ scaling is mapped into the $\ln(V_{S30})$ scaling due to the correlation between
 132 $\ln(Z_{2.5})$ and $\ln(V_{S30})$ in the empirical data.

133 Using the 3-D simulations for a single V_{S30} , the scaling for $Z_{2.5}$ is modeled by:

$$f_{site}(Z_{2.5}) = c_0 + c_2 \ln(Z_{2.5}) \quad (10)$$

134 This basin scaling from the simulations can be added to the site term in equation (9) as the
 135 slope of the normalized basin depth:

$$\begin{aligned} f_{site}(Z_{2.5}, V_{S30}) &= c_0 + c'_1 \ln(V_{S30}) + c_2 \ln(Z'_{2.5}) \\ &= c_0 + c'_1 \ln(V_{S30}) + c_2 \ln(Z_{2.5}) - c_2 (\ln(Z_{2.5ref}(V_{S30}))) \\ &= c_0 + c'_1 \ln(V_{S30}) + c_2 \ln(Z_{2.5}) - c_2 (a_1 + a_2 \ln(V_{S30})) \quad (11) \\ &= (c_0 - c_2 a_2) + (c'_1 - c_2 a_2) \ln(V_{S30}) + c_2 \ln(Z_{2.5}) \\ &= (c_0 - c_2 a_2) + c_1 \ln(V_{S30}) + c_2 \ln(Z_{2.5}) \end{aligned}$$

136 The last line of equation (11) shows that by using the normalized basin in the GMM, we
 137 can recover the correct scaling with both V_{S30} and $Z_{2.5}$. This is the approach used in this
 138 study.

139 We emphasize that the purpose of using the normalized basin depth in the GMM is not
 140 to impose the correlation observed in the empirical data set to all other sites but rather to
 141 remove the effects of the correlation in the empirical data set from the estimation of the site
 142 terms at new sites for which this correlation may not apply.

143 In practice, the $Z_{2.5}$ value for a new site may not be known. In this case, the default $Z_{2.5}$

given the V_{S30} is often used. If the default basin depth is used, then there is an assumption that the correlation between V_{S30} and $Z_{2.5}$ in the empirical data set applies to the new site; however, this correlation may not be applicable to the region of interest. A better approach is to develop an appropriate relation between V_{S30} and $Z_{2.5}$ for the region of interest and use that relation to define the $Z_{2.5}$ value for sites without a measured $Z_{2.5}$.

3-D SIMULATED GROUND-MOTION DATA SET

Frankel et al. (2018) and Wirth et al. (2018) generated a set of broadband (0–10 Hz) simulated seismograms for M9 Cascadia interface events based on the combination of 3-D finite-difference simulations (≤ 1 Hz) and 1D stochastic synthetics (≥ 1 Hz). The source model used in the simulations was informed by observations of 2011 Tohoku, Japan M9 earthquake and the 2010 Maule, Chile M8.8 earthquake. For sites not in sedimentary basins, the response spectral values from the simulation results are consistent with the Abrahamson et al. (2016) GMM (called BCHydro2016) for periods between 0.1 to 6.0 sec. The finite-difference simulated and the stochastic synthetics are combined using matched filters at 1 Hz.

The broadband simulated seismograms and the response spectra for 30 realizations of M9 earthquakes on the Cascadia subduction zone are described by Frankel et al. (2018) and Wirth et al. (2018). The 30 realizations include different hypocenters, slip distributions, and positions of the eastern edge of the rupture. The kinematic rupture model consists of high stress-drop M8 subevents on the deeper portion of the rupture zone combined with shallower background slip on the shallower part of the rupture zone. A map with the epicenters of the 30 simulations is shown in Figure 3. The epicenters and the down-dip edge for the 30 events are listed in Table 1.

The 3-D velocity model used the P- and S-wave velocity model developed by Stephenson et al. (2017) (a cutaway view in Fig. 1 of Frankel et al., 2018). The 3-D model extends from Cape Mendocino in the south to the middle of Vancouver Island in the north and includes

the subducted slab and the plate interface (McCrory et al., 2012). The deep basins, such as the Seattle and Tacoma basins, reach depths of 7-8 kilometers. The tomography results for the Seismic Hazards Investigation in Puget Sound (SHIPS) project and the crustal velocities of Moschetti et al. (2010) are also considering in this model. The minimum S-wave velocity was set at 600 m/s, which is a typical shear-wave velocity for the surficial glacial sediments in Seattle. To account for realistic small-scale variations in V_S , a random component was added to the V_S in the Quaternary sediments in the top 1.3 km of the model. Figure 4 shows the T=3 sec PSA map of the Puget Lowland area for one simulation scenario (ID1, epicenter 43.617°N, 125.417°W) from Frankel et al. (2018). The large site amplifications are located in the Seattle and Tacoma basins, and the overall amplification of the Puget Lowland is higher than for the outside area.

The velocity model covers approximately 40.2°N to 50°N and 121°W to 129°W with $Z_{2/5}$ values from about 100 to 7000 m depth (Stephenson et al. (2017)). In this study, we limited the region for the non-ergodic GMM development to the Puget Sound region: 46.9°N to 48.3°N, and 121.5°W to 123.2°W. The spatial distribution of basin-depth parameter ($Z_{2.5}$) in the study region is shown in Figure 5. For most of the region, the $Z_{2.5}$ is greater than 1000 m. A histogram of the sampling of the $Z_{2.5}$ values is shown in Figure 6. With this sampling, the simulations provide constraints for a wide range of $Z_{2.5}$ values.

MODIFIED BASIN-DEPTH SCALING

As a first step, the basin scaling in the AG20 model is modified to be consistent with the basin scaling in the M9 project simulations. To evaluate the basin scaling from the simulations, the total residuals, δ , from the 3-D simulations are computed relative to the AG20 model without the basin term:

$$\delta = \ln(PSA) - ((AG20(M, R, \dots) - f_{basin}(Z_{2.5})) \quad (12)$$

The total residuals at six periods (0.2, 1.0, 2, 3, 5, and 10 sec) are shown as a function of $Z'_{2.5}$ in Figure 7. The squares show the mean residual for different $Z'_{2.5}$ bins. The mean residual for each $Z'_{2.5}$ bin and the corresponding range of $Z_{2.5}$ values are listed in Table 2.

For the long periods ($T \geq 1$), the residuals are approximately constant at the smaller $Z'_{2.5}$ values ($Z'_{2.5} < 0.3$), whereas there is a near-linear dependence with $\ln(Z'_{2.5})$ for larger $Z'_{2.5}$ values. Because a uniform V_{S30} of 600 m/s is used in the simulations, $Z_{ref}=2000$ m for all sites, and $Z'_{2.5}=0.3$ corresponds to $Z_{2.5}=560$ m. In contrast, the residuals at short periods (e.g. $T = 0.2$ sec) do not show a trend with basin depth because the stochastic model used for frequencies greater than 1 Hz used a 1-D velocity structure rather than the 3-D velocity structure.

The basin-depth term in the AG20 model is shown by the blue curve in Figure 7. For the deep basin sites, the slopes of the residuals are similar to the basin-depth scaling from the AG20 model. A key difference between the basin scaling in the AG20 model and the basin scaling from the 3-D simulations is the extension of the scaling to $Z'_{2.5}$ less than 1 ($Z_{2.5} < 2000$ m) seen in the simulation results.

We modified the basin-depth scaling of the AG20 model to be consistent with the basin scaling from the M9 simulations. The total residuals were fit to the following basin-depth model:

$$f_{basin3d}(Z_{2.5}) = \begin{cases} b_1 \ln(Z'_{2.5}/Z'_x) + b_2 & \text{for } Z'_{2.5} > Z'_x \\ b_2 & \text{for } Z'_{2.5} \leq Z'_x \end{cases} \quad (13)$$

An example of the new basin-depth scaling for $T=3$ sec is shown in Figure 8. The red solid curve is the fit the residual and the blue curve is the AG20 basin scaling (not fit to the simulation residuals). For $T=3$ sec, the M9 simulations have $Z'_{2.5}$ slope similar to the AG20 basin scaling for deep basin sites ($Z'_{2.5} > 1$ or $Z_{2.5} > 2000$ m) even though these two models are based on different ray paths: the empirical AG20 basin terms are based on ray paths from intra-slab events, whereas the M9 simulations are based on ray

paths from interface events. There is an offset between the fit to the residuals and the AG20 model. Because the basin scaling in the AG20 model is normalized to be zero at $Z'_{2.5}=1$, the difference between the two curves at $Z'_{2.5}=1$ reflects the average scale factor between the 3-D simulations and the AG20 model. We interpret this difference as the average difference in the source as modeled by the 3-D simulations and a represented in the AG20 GMM.

The objective of this study is to modify the AG20 GMM to capture the non-ergodic effects without changing the overall level of ground motion. The median $Z_{2.5}$ for a given V_{S30} in the AG20 data set corresponds to $Z'_{2.5}=1$. Therefore, we used the difference between the fit to the residuals and the AG20 basin term at $Z'_{2.5}=1$ as the constant shift to center the simulations on the AG20 model. Using this shift allows us to incorporate the basin scaling from the 3-D simulations into the AG20 GMM without changing the average ground-motion level given by the AG20 GMM.

The new basin-depth term is compared with the AG20's basin-depth term for six periods in Figure 9. The period dependence of the basin-depth coefficients from the M9 simulations (b_1) and from the AG20 model (a_{39} term) are compared in Figure 11a. The $Z'_{2.5}$ slopes from the M9 simulations are slightly steeper than the $Z'_{2.5}$ slopes in the AG20 model for periods between 2 sec to 5 sec. At other periods, the slopes for the $Z'_{2.5}$ scaling in the simulations are flatter than in the AG20 model.

The depth at which the basin scaling goes flat is shown in Figure 11c: it increases from $Z_{2.5} = 170$ m ($Z_{2.5'} = 0.106$) at $T=1$ sec to $Z_{2.5} = 930$ m ($Z_{2.5'} = 0.477$) at $T= 10$ sec, whereas in the AG20 model, Z_x is fixed to 1.0 corresponding to 5000 m for $V_{S30}= 200$ m/s and 2000 m for $V_{S30} > 570$ m/s.

The constant shift (C_{SIM}) between the simulations and the AG20 model at $Z'_{2.5}=1$ is shown in Figure 11b. On average, the simulated $\ln(PSA)$ values are about 0.7 for the period range of 1.5 to 3 sec, corresponding to a factor of 2 increase over the AG20 model for this period range. While we do not use this difference in our modified AG20 model for non-ergodic effects, these terms can be used to scale the AG20 model if a model consistent

with the overall level of ground motion from the M9 simulations is desired. The coefficients of equation (13) for 1 sec to 10 sec are listed in Table 3.

The updated basin scaling is also shown in terms of the $Z_{2.5}$ without the normalization in Figure 10. This is the same information as shown in Figure 9, but is easier to interpret without the normalization.

Maps of the basin terms for $T=3$ sec for the original AG20 model and the modified AG20 model are shown in Figure 12. For this period, the main difference is for sites with $Z_{2.5} < 2000$ m, with the modified basin scaling leading to reduced ground motions for the shallow $Z_{2.5}$ values as shown by the light-blue and dark-blue regions.

The AG20 model with the modified basin-depth scaling is given by:

$$AG20_{SIM-erg}(M, R, \dots) = AG20(M, R, \dots) - f_{basin}(Z_{2.5}) + f_{basin3d}(Z_{2.5}) - C_{SIM} \quad (14)$$

In the remainder of the paper, we refer to this modified basin-term model as $AG20_{SIM-erg}$. It is called an "ergodic" model for Cascadia because it applies to the entire region even though it is not a global model and includes region-specific terms. This simplifies the terminology when we later incorporate the partially non-ergodic site terms into the modified AG20 GMM for Cascadia.

Extrapolation to Short Periods

The M9 Project simulations combine 3-D simulations for low frequencies with 1-D stochastic simulations for high frequencies. The two sets of simulations are combined using match filters at 1 Hz to give broadband simulations. As a result, the basin effects are not represented at high frequencies and are only partially represented at the 1 Hz match-filter frequency.

The basin scaling from the simulations is compared to the basin scaling from the AG20 model in Figure 13. The simulations show a rapid decrease in the basin term for periods

less than 2 sec, reflecting the match filters. The empirically based basin scaling in the AG20 model decays more slowly for periods less than 2 sec than for the simulations, reaching zero at $T=0.2$ sec. We assume that the basin effects in the simulations are fully captured for periods of 2 seconds or greater, but that they are underestimated for periods less than 2 sec.

The basin scaling at long periods from the simulations is extrapolated to periods less than 2 sec using linear interpolation between 0.2 sec and 2 sec on the log-period axis. To have a smooth transition at $T=0.2$ sec, a cubic function is used with constraints that the value and the first derivative are both zero at $T=0.2$ and the derivative equals the slope of the linear interpolation at the midpoint (in log units) between 0.2 and 2 sec. The taper function is zero at $T = T_1$ and unity at $T = T_2$. The taper function is given by:

$$Taper_1(T) = \begin{cases} 0 & \text{for } T < T_1 \\ 4 \left(\frac{\ln(T/T_1)}{\ln(T_2/T_1)} \right)^2 - 4 \left(\frac{\ln(T/T_1)}{\ln(T_2/T_1)} \right)^3 & \text{for } T_1 < T < \sqrt{T_1 T_2} \\ \frac{\ln(T/T_1)}{\ln(T_2/T_1)} & \text{for } \sqrt{T_1 T_2} < T < T_2 \\ 1 & \text{for } T > T_2 \end{cases} \quad (15)$$

In this application, $T_1=0.2$ sec and $T_2=2$ sec.

The extrapolation to shorter periods scales the basin term for $T=2$ sec from the 3-D simulations by the taper. With this extrapolation, the basin term is given by:

$$f_{basin3d}(T) = \begin{cases} f_{basin3d}(T=2)Taper_1(T) & \text{for } T < 2sec \\ f_{basin3d}(T) & \text{for } T \geq 2sec \end{cases} \quad (16)$$

This extrapolation is shown by the short dashed lines in Figure 13.

Using these basin factors, the median response spectra for an M9 interface earthquake at a distance of 100 km and $V_{S30}=600$ m/s are shown for $Z_{2.5}$ values between 100 m and 8000 m in Figure 14. For $Z_{2.5}$ values less than 1000 m, there is strong flattening in the spectral shape for periods greater than 3 sec. This change in spectra shape for sites with low $Z_{2.5}$

values is seen in the 3-D simulations, but it is an unusual shape for empirical GMMs. This shape can be removed by using a lower limit of 1000 m for $Z_{2.5}$, but for this study, we follow the trends of the simulations and do not apply a lower limit to the $Z_{2.5}$ values.

Comparison with SDCI Basin Factors

The models for the basin scaling from the AG20 GMM (f_{basin}) and the 3-D simulations ($f_{basin3d}$) are compared to the SDCI basin factors in Figure 15. For both the f_{basin} and $f_{basin3d}$ models, the basin factors are normalized to a reference $Z_{2.5}$ of 3000 m to be consistent with the SDCI definition. The SDCI basin factors approximately envelop the normalized basin factors from the empirical AG20 model and from the 3-D simulations for $Z_{2.5}$ less than 7000 m.

Evaluation of Residuals for the Modified AG20 model

Using the $AG20_{SIM-erg}$ model that includes the modified basin-depth term, we recomputed the residuals for the M9 simulations using the following model:

$$\ln(PSA) = AG20_{SIM-erg}(M, R, ...) + C_{SIM} + \delta B + \delta S2S + \delta WS \quad (17)$$

in which δB is the between-event residual, $\delta S2S$ is the between-site residual, and δWS is the within-site residual. The total residuals for $T = 2$, $T=3$, and $T=5$ sec are shown as functions of R_{RUP} and $Z_{2.5}$ in Figure 16. The similar plots for the within-site residuals are shown in 17. For distances up to 250 km covered by the simulations, the distance scaling of the $AG20_{SIM-erg}$ model is consistent with the distance scaling of the M9 simulations even though the regionalized distance terms in the AG20 model are based on intra-slab data, whereas the distance scaling in the M9 simulation is for interface events. The M9 simulation results are limited to rupture distances less than 250 km, so this comparison does not provide a check on the distance attenuation in the AG20 model for interface events at larger distances

covering eastern Washington and Idaho.

Although the average distance scaling from the M9 simulations is consistent with the AG20 distance scaling, two realizations have different distance scaling from the AG20 GMM: event ID 9 (43.75°N, 123.98°W) has steeper attenuation with distance and event ID 22 (46.77°N, 125.44°W) has flatter attenuation with distance as shown in Figure 18. Event ID 9 has the eastern-most down-dip edge of the source model and event ID 22 has the western-most down-dip edge. Other events with similar epicenter locations and with the same down-dip edges do not show similar differences in the distance scaling, so this does not appear to be a systematic effect due to epicenter location or the down-dip edge location.

The between-event residuals are shown as a function of period in Figure 19. Two events (ID 4: 48.25°N, 124.67°W and ID 12: 46.87°N, 124.54°W) have much more negative event terms than the other 28 events and appear as outliers. Both events ID 4 and ID 12 have epicenters located in the northern end of the rupture and both are for the middle down-dip edge. The other events with epicenters to the north and with the middle down-dip edge (ID10, ID 15 and ID 16) do not have a large negative event term. Excluding the two outliers would increase the median from the simulations by 5-10%, but it would not affect the non-ergodic GMM.

Based on this evaluation of the residuals, we conclude that the distance scaling in the $AG20_{SIM-erg}$ model is consistent with the distance scaling in the simulations for the distances up to 250 km. Therefore, no modifications are made to the distance scaling in the AG20 GMM.

In the following model development, we used the M9 data from all 30 events to modify the AG20 GMM to be a partially non-ergodic GMM, including the four events discussed above that appear as outliers in amplitude or distance scaling.

NON-ERGODIC SITE TERMS

Following the notation of Al Atik et al. (2010), the general form for the median for a non-ergodic GMM can be written as:

$$\mu_{nonerg}(M, R, S, \dots, t_e, t_s) = \mu_{erg}(M, R, S, \dots) + \delta L2L(t_e) + \delta S2S(t_s) + \delta P2P(t_e, t_s). \quad (18)$$

in which $\mu_{erg}(M, R, S, \dots)$ is the ergodic GMM model, $\delta L2L$ is the median shift in the source term for the earthquake location, $\delta P2P$ is the median shift in the path term (distance scaling) for the earthquake and site location, and t_e and t_s are the coordinates of the earthquake and the site, respectively. The $\delta S2S$ term is the same as in equation 17. The fully non-ergodic GMMs account for the systematic source and path effects in addition to the systematic site effect by including the non-ergodic terms that depend on the geographical locations of the source and/or site.

In the M9 simulations, the sources are all M9 Cascadia earthquakes, so we cannot separate the path effects from the site effects. Therefore, we develop a non-ergodic Cascadia GMM that combines the path and site terms into a single site term and uses the AG20 model with the modified basin-depth scaling as the reference ergodic model. The median is given by:

$$\mu_{nonerg-CAS}(M, R, S, \dots, t_s) = AG20_{SIM-erg}(M, R, \dots) + \delta S2S(t_s). \quad (19)$$

The between-site residuals, $\delta S2S$, are shown as functions of distance and $Z_{2.5}$ in Figure 20. Again, there are no significant trends in the total residuals with distance or $Z_{2.5}$; however, there is larger variability for $Z_{2.5} > 1000$ m reflecting greater variability for sites in basins than for sites outside the basins.

Non-Ergodic GMM Methodology

We could just treat the $\delta S2S$ in equation (19) as a site-specific constant that is independent between sites. Instead, we used the non-ergodic GMM framework that includes the spatial correlation of the non-ergodic terms. This leads to a smoother spatial variation of the non-ergodic site terms than the traditional random-effects approach. An added advantage of using the non-ergodic framework is that it allows us to test the applicability of the non-ergodic GMM methodology for very large data sets. As we move to including non-ergodic path effects in the future, this more complex methodology will need to be used in place of simple random-effects regression.

In recent studies (e.g. Landwehr et al., 2016; Lavrentiadis et al., 2021; Meng and Goulet, 2021; Sung et al., 2021), the non-ergodic terms were estimated using the variable coefficient model (VCM) approach which is based on Gaussian Process regression with a hierarchical Bayesian framework using Monte-Carlo Markov-Chain (MCMC) sampling as given in the program Rstan or STAN (Stan Development Team, 2020); however, MCMC becomes computationally expensive for a very large data set. For the M9 project simulations, there are 30 realizations for over 20,000 sites, leading to over 600,000 ground-motion values which makes MCMC not practical. Therefore, we adopted the Integrated Nested Laplace Approximation (INLA) approach (Rue et al., 2009, 2017) as an alternative approach that is computationally faster with reduced memory requirements. INLA (Lindgren et al., 2015) focuses on individual posterior marginal distributions rather than estimating the joint posterior distribution of the model parameters and can be expressed as latent Gaussian Markov random fields (GMRF). The latent GRMF structure of the model will have zero mean and precision matrix Q (Σ^{-1}), which is based on a vector of hyperparameters (θ) and which will often be very sparse, reducing the computation time.

The full joint distribution of the observations (\mathbf{y}), the latent process (\mathbf{x}) and the hyper-

parameters (θ) can be written as follows:

$$p(\mathbf{x}, \theta | \mathbf{y}) = \frac{p(\mathbf{y} | \mathbf{x}, \theta) p(\mathbf{x}, \theta)}{p(\mathbf{y})} = p(\mathbf{y} | \mathbf{x}, \theta) p(\mathbf{x} | \theta) p(\theta) \quad (20)$$

in which \mathbf{y} is a vector of n observations $= (y_1, \dots, y_n)$; $p(\mathbf{y})$ is the marginal likelihood, which is a normalizing constant and is often ignored. $p(\mathbf{x}, \theta)$ is the joint distribution for the random effects and the hyperparameters $= p(\mathbf{x} | \theta) p(\theta)$. $p(\theta)$ is the prior distribution of the ensemble of hyperparameters.

Moreover, the statistical framework for Bayesian inference in the previous models can be further written in the following forms:

$$\log p(y | x, \theta) = \sum_i \log p(y_i | x_i, \theta) \quad (21)$$

$$\log p(x | \theta) \propto \frac{1}{2} \log |Q(\theta)| - \frac{1}{2} (x^T Q(\theta) x) \quad (22)$$

$$\theta \sim p(\theta) \quad (23)$$

Here, the distribution of the observation y at index i depends only on the value of the latent variable x_i for the Gauss Markov random field. x_i is a multivariate normal distribution with a sparse precision matrix Q . The precision matrix has mostly zeroes (e.g., x_i and x_j is the conditionally independent), and the zeroes are not be stored in memory which reduces the memory requirements

Within the INLA framework, we do not attempt to estimate the full joint posterior distribution, $p(x, \theta | y)$, but rather estimate the univariate posterior marginals of the latent effects, $p(x | y)$ and hyperparameters, $p(\theta | y)$. Rue et al. (2009) provides a method for approximating the marginal distributions using different approximations (Gaussian, Laplace, simplified Laplace) to the distributions involved and adopting the Laplace approximation for the inte-

grals to obtain the approximation for $p(x|y)$ and $p(\theta|y)$ per element.

The approximation for the latent effects is given by:

$$p(x_i|y) = \int p(x_i|\theta, y) p(\theta|y) d\theta \simeq \sum_k \tilde{p}(x_i|\theta_k, y) \tilde{p}(\theta_k|y) \Delta_k = \tilde{p}(x_i|y) \quad (24)$$

$$\tilde{p}(x_i|\theta, y) = \mathcal{N}\{x_i, \mu_i(\theta), \sigma_i^2(\theta)\} \quad (25)$$

The approximation for hyperparameters is given by:

$$p(\theta_k|y) = \int p(\theta|y) d\theta_{-k} \simeq \int \tilde{p}(\theta|y) d\theta_{-k} = \tilde{p}(\theta_k|y) \quad (26)$$

in which $\tilde{p}(\cdot|\cdot)$ is an approximated density of its arguments; $\mu(\theta)$ is a vector of the mean of the approximation; $\sigma^2(\theta)$ is a vector of corresponding marginal variances; Δ_k is the weight associated with θ_k for the hyperparameters in the element; θ_{-k} is a vector of hyperparameter θ without the element θ_k .

For the R-INLA package, there are tools that can sample from the posterior distribution of the model at the mesh nodes and interpolate the distribution to new locations. The mean (μ_i) and variance (σ_i^2) of the marginal distribution for the individual latent effect x_i (non-ergodic term) are given by:

$$E[x_i] = \mu_i = \sum_j \omega_j \mu_j \quad (27)$$

$$Var[x_i] = \sigma_i^2 = \sum_j \omega_j (\mu_j - \mu)^2 \quad (28)$$

in which ω_j is the weight, μ_j is the j th sample of the marginal distribution.

The non-ergodic site term were estimated using the computer software R-ILAN based

on the Matern covariance function. The covariance function between a pair of sites for $cov(x_{sta}, x'_{sta})$ is defined as

$$cov(x_{sta}, x'_{sta}) = \Theta^2 \frac{2^{1-v}}{\Gamma(v)} (\sqrt{8v} \|x_{sta} - x'_{sta}\|/\rho)^v K_v(\sqrt{8v} \|x_{sta} - x'_{sta}\|/\rho). \quad (29)$$

in which Γ is the Gamma function; ρ is the correlation length, Θ is the marginal standard deviation of the spatial field; v is the smoothness parameter; x_{sta} and x'_{sta} are the coordinates of the two sites depending on the coefficient; K_v is the modified Bessel function of second kind and order $v > 0$.

The resulting hyperparameters of equation (29) are listed in Table 4 and shown in Figure 21. The correlation lengths of simulation are about 40 km for all periods which is greater than the correlation length of empirical data of California (~ 10 km) found by Lavrentiadis et al. (2021). This larger correlation length from the M9 simulations compared to empirical data for California earthquakes may be due to the single path in the simulations that does not capture the variability from different ray paths. Variable path effects are expected to reduce the spatial correlation lengths of the path effects.

Maps of the non-ergodic site term, $\delta S2S$, for $T=3$ sec and $T=5$ sec are shown in Figure 22. These site terms show the larger basin terms (positive values) for sites located close to the western portion of the Seattle basin and the eastern portion of the Tacoma basin.

For comparison, we also computed the site term using a standard random-effects approach (without spatial correlation of the $\delta S2S$). These maps are shown in Figure 23. We then computed the difference between the site terms based on the random-effects without spatial correlation ($\delta S2S$) and the non-ergodic site terms with spatial correlation ($\delta S2S_{nonerg}$). The spatial distributions of the differences for $T=3$ and $T=5$ sec are shown in Figure 24. The largest differences are at basin edge sites where there are rapid changes in the basin terms. The differences are shown as a function of $Z_{2.5}$ in Figure 25. The mean difference is near zero (no bias) and the standard deviation of the difference is less than 0.052 natural log units,

but the differences are large at a few sites. For $T=3$ sec, the largest differences are ± 0.25 natural log units. For all periods, the largest differences are for $T=2$ sec at basin sites with $Z_{2.5} > 1000$ m, with differences up to ± 0.5 natural log units. Figure 26 shows the map of differences for $T=2$ sec for a cross section that goes through the maximum differences.

The non-ergodic site term is the deviation from the average basin scaling given by the GMM. The total basin term is the sum of the new basin-depth scaling and non-ergodic site term. The map of the total basin term for $T=3$ sec is shown in Figure 27. By including the non-ergodic site terms, the large site terms are not only in the deepest part of the basin. For example, for the Seattle basin, there is a broad zone of large amplification that extends west of the deepest part of the basin.

ALEATORY VARIABILITY

The total standard deviation (σ) is computed from the sum of the between-event variance (τ^2) and the within-event variance (ϕ^2):

$$\sigma = \sqrt{\tau^2 + \phi^2} \quad (30)$$

The within-event standard deviation is decomposed into the between-site standard deviation, ϕ_{S2S} , and the within-site (single-station) standard deviation, ϕ_{SS} .

$$\phi = \sqrt{\phi_{S2S}^2 + \phi_{SS}^2} \quad (31)$$

The single-station total standard deviation, σ_{SS} , is given by

$$\sigma_{SS} = \sqrt{\tau^2 + \phi_{SS}^2} \quad (32)$$

For seismic hazard calculations, σ is used as the aleatory variability for the ergodic GMM, and σ_{SS} is used for the aleatory variability for the partially non-ergodic GMM with

the site-specific site effects removed. The ϕ_{S2S} represents the epistemic uncertainty in the non-ergodic site term for sites well outside the region with data (e.g., outside the region by more than the correlation length).

The sigma models for the AG20 model are also included in our comparisons. The AG20 global τ model and the AG20 Group 1 ϕ , ϕ_{SS} , and ϕ_{S2S} models that are recommended for application to Cascadia are given below:

$$\tau = 0.47 \quad (33)$$

$$\phi = \phi_1 \quad (34)$$

$$\phi_1^2(R_{RUP}, T) = \begin{cases} d_1(T) & \text{for } R_{RUP} < 150 \\ d_1(T) + d_2(T) \left(\frac{R_{RUP} - 150}{300} \right) & \text{for } 150 \leq R_{RUP} \leq 450 \\ d_1(T) + d_2(T) & \text{for } R_{RUP} > 450 \end{cases} \quad (35)$$

$$\phi_{S2S}(T) = \sqrt{\phi_1^2(R_{RUP} = 100, T) - 0.165} \quad (36)$$

$$\phi_{SS}(R_{RUP}, T) = \sqrt{\phi^2(R_{RUP}, T) - \phi_{S2S}^2(T)} \quad (37)$$

The d_1 and d_2 coefficients for the ϕ_1^2 model are listed in Table 5

Aleatory Variability Models

The four components of the aleatory variability, τ , ϕ , ϕ_{SS} , and ϕ_{S2S} , are shown in Figure 28. For each standard deviation component, the results from the simulations are compared to the AG20 model for Cascadia given in equations (33) to (37). The values of ϕ , ϕ_{SS} , and ϕ_{S2S} for the 3-D simulations and the AG20 GMM are listed in Table 6.

Tau

For τ (Figure 28a), the between-event variability from the simulations is less than the between-event variability from the AG20 model. The AG20 τ model is a global value from interface and intra-slab earthquakes. The τ from the simulations is similar to the AG20 τ for periods greater than 2 sec which is the deterministic part of the simulations. For shorter periods, the stochastic simulations lead to much smaller τ values than the AG20 model. In PSHA applications, we recommend using the global τ from the AG20 GMM because it is empirically based and does not have to very small τ values for periods less than 2 sec.

Phi

For ϕ (Figure 28b), the within-event variability from the simulations relative to the AG20 model (green curve) is large for $T=1.5$ to $T=4$ sec. This is due to the limit on the basin-depth scaling to deep basin sites ($Z'_{2.5} > 1$) in the AG20 model. With the modified basin-depth scaling, the ϕ from the simulations (blue curve) is similar to the Cascadia ϕ model from AG20 (black curve) for $T=1.5$ to $T=4$ sec. The simulations lead to smaller ϕ values for long periods ($T>5$ sec). For comparison, the ϕ_{SS} from the simulations is shown by the red curve. The strong period dependence seen in the ϕ values is not seen in the ϕ_{SS} values, indicating that the period dependence is related to the site terms.

The distance dependence of ϕ is shown in Figure 29 for periods of 0.2, 1, 2, 3, 5, and 10 sec. For distances of 50-120 km, the ϕ for the $AG20_{SIM-erg}$ model is approximately constant. There is a trend of decreasing ϕ with increasing distance for distances of 120-200 km. The ϕ model from AG20 (black curve) is slightly lower than the ϕ from the simulations for periods of 2-3 sec, and it is larger than the ϕ from the simulations at longer periods. This indicates that the 3-D basin effects for periods of 2-3 seconds are more variable in the Seattle region than in other subduction zone regions, whereas at 5-10 sec, the other regions show more variability.

PhiSS

For ϕ_{SS} (Figure 28c), the standard deviations are similar for all the models based on the 3-D simulations because the site term is removed. At long periods ($T > 1$ sec), the ϕ_{SS} for the simulations is smaller than the global ϕ_{SS} for the AG20 model: the ϕ_{SS} for the simulations is about 0.3 compared to 0.41 for the AG20 model.

The distance dependence of ϕ_{SS} is also shown in Figure 29 (red curve) for periods of 0.2, 1, 2, 3, 5, and 10 sec. There is a weaker distance dependence (120-200 km) for ϕ_{SS} than for ϕ .

PhiS2S

For ϕ_{S2S} (Figure 28d), the between-site variability from the simulations relative to the AG20 model (green curve) is large for $T=1.5$ to $T=4$ sec. As discussed for the ϕ model, this increase is due to the limit on the basin-depth scaling to deep basin sites ($Z'_{2.5} > 1$) in the AG20 model. The ϕ_{S2S} values estimated with and without spatial correlation are similar to each other due to the dense sampling of the simulations. The ϕ_{S2S} values from the simulations are similar to the ϕ_{S2S} model from AG20 at periods between 1.5 and 5 sec. This indicates that the variability in the site effects due to the 3-D velocity model for Cascadia is similar to the 3-D effects for sites in other regions represented in the empirical data used by AG20 for $T=1.5$ sec to $T=5$ sec.

The dependence of the ϕ_{S2S} on $Z_{2.5}$ is shown in Figure 30 for periods of 1, 2, 3, and 5 sec. Again, the larger ϕ_{S2S} values for the AG20 model in the 2-5 sec period range for $Z_{2.5}$ values of 1000 to 2000 m are due to the lack of scaling for low $Z_{2.5}$ values in the AG20 GMM. There is an increase in the ϕ_{S2S} for $Z_{2.5}$ values between 1000 and 3000m indicating that the basin factors for specific sites are more variable over this range of depth depths and may reflect basin edge effects.

The ϕ_{S2S} values from the AG20 model range from 0.3 to 0.4 and are based on data from three regions (Cascadia, New Zealand, and Taiwan). This model did not consider a

$Z_{2.5}$ dependence because $Z_{2.5}$ values were not available for regions other than Cascadia and Japan.

Abrahamson and Gulerce (2020) found a strong regional dependence for ϕ_{S2S} . Therefore, a Cascadia-specific ϕ_{S2S} is developed based on the 3-D simulations and the AG20 ϕ_{S2S} model. At short periods ($T \leq 1.0$ sec), the M9 simulations do not capture the full variability, so the ϕ_{S2S} values from the AG20 model are assumed to apply to Cascadia. At intermediate periods ($1.5 \leq T \leq 5$), the ϕ_{S2S} from the 3-D simulations are slightly larger than the ϕ_{S2S} from the AG20 model. The results from the 3-D simulations are considered reliable for this period range and are used for the Cascadia-specific model. At long periods ($5 < T < 10$ sec), the ϕ_{S2S} from the 3-D simulations shows a strong reduction as the period increases. This feature was also seen in the empirical data in the AG20 model development, but a more gradual decrease was imposed on the model to avoid a large change in the shape of the uniform hazard spectra at long return periods. For the Cascadia-specific ϕ_{S2S} , we used the average of the ϕ_{S2S} from the 3-D simulations and the ϕ_{S2S} from the AG20 model to smooth out the period dependence. The resulting Cascadia-specific ϕ_{S2S} is given in the last column of Table 6.

Total Sigma

The total standard deviations for the AG20 model, $AG20_{SIM-erg}$ model, and AG20 non-ergodic model are compared in Figure 31. The sigma values are shown for three combinations of the ϕ and τ : the first set uses the ϕ and τ from the simulations; the second set uses the empirical ϕ and τ from the AG20 GMM; and the third set uses the ϕ from the simulations and the global τ from the AG20 GMM. Table 7 lists the digital values for the total standard deviation for the different models.

The total standard deviation computed using the ϕ for the $AG20_{SIM-erg}$ model with the global τ (dashed-blue curve) is similar to the empirical AG20 model (black curve) for periods between 1.5 and 4 sec. For the $AG20_{SIM-erg}$ model and the non-ergodic model, there is a

slight reduction in σ as the period increases from 2-10 sec, consistent with the trend in the empirical AG20 sigma model.

Single-Station Sigma

The period dependences of the σ_{SS} models are shown in Figure 32a. The σ_{SS} from the simulations ranges from 0.50 to 0.58. In comparison, the AG20 σ_{SS} model for Cascadia is a period-independent 0.62. For reference, the AG20 σ model for Cascadia is also shown in Figure 32a by the solid black curve.

Studies of crustal ground-motion data have found that the single-station sigma is typically 15-20% smaller than the ergodic sigma for long-period ground motions in California (Lin et al. (2011)). The ratios of the single-station sigma to total sigma are shown in Figure 32b. The ratio for the AG20 model is between 0.84 and 0.90. For the simulations, there is a greater reduction with ratios between 0.75 and 0.85.

In most applications of partially non-ergodic GMMs, a single average σ_{SS} model is used for all sites in the region. This is like an ergodic assumption for the variability term. With the 30 events for the 3-D simulation, the single-station sigma can be computed for each site. These site-specific standard deviations, denoted $\sigma_{SS,s}$, are shown on a map for $T = 1.5$ sec and 3 sec on Figure 33 and Figure 34. These maps show that the $\sigma_{SS,s}$ is larger near the edges of basins.

The site-specific single-station sigma is also computed for different distance bins as shown in Figure 35 for periods of 0.2 sec to 10 sec. For periods of 1.5 - 10 sec, there is no significant distance dependence of the $\sigma_{SS,s}$ for distances less than 200 km. For $T=0.2$ sec and $T=1$ sec, there is a trend of decreasing σ_{SS} with increasing distance, but the σ_{SS} values for the simulation for these short periods are much lower than for the longer periods.

NON-ERGODIC PSHA IMPLEMENTATION

For traditional ergodic GMMs, the seismic hazard for the i^{th} source is given by:

$$Haz_i(Y > z) = N_i(M_{min}) \int_M \int_R f_{m_i}(M) f_{R_i}(R, M) P(Y > z | M, R, S) dR dM \quad (38)$$

in which $N_i(M_{min})$ is the rate of earthquake with $M > M_{min}$, $f_{m_i}(M)$ is the magnitude probability density function, and $f_{R_i}(R, M)$ is the distribution of distances for a given magnitude for the i^{th} source. The conditional probability of exceeding the ground motion for a given scenario is given by:

$$P(Y > z | M, R, S) = 1 - \Phi \left(\frac{\ln(z) - \mu_{erg}(M, R, S)}{\sigma_{erg}(M, R)} \right) \quad (39)$$

in which $\mu_{erg}(M, R, S)$ is the median ground motion in \ln units, $\sigma_{erg}(M, R, T)$ is the standard deviation, and $\Phi()$ is the cumulative distribution function of a standard normal distribution. For a partially non-ergodic GMM, the $P(Y > z | M, R, S)$ term is modified as follows:

$$P(Y > z | M, R) = 1 - \Phi \left(\frac{\ln(z) - [\mu_{erg}(M, R, S) + \delta S 2 S_{nonerg}(Lat_s, Lon_s)]}{\sigma_{SS}(M, R, Lat_s, Lon_s)} \right) \quad (40)$$

in which Lat_s and Lon_s are the coordinates of the site locations; σ_{SS} is the single-station aleatory standard deviation, which can be a regional value or a site-specific value.

Implementing a partially non-ergodic GMM in probabilistic seismic hazard requires three changes to the ergodic GMM:

- Modify the the central estimate of the median ground motion by the non-ergodic site term, $\delta S 2 S_{nonerg}$
- Reduce the aleatory variability to the σ_{SS} value or to the site-specific $\sigma_{SS,s}$ value
- Include the epistemic uncertainty in the value of the non-ergodic site term in the GMM

579 logic tree

580 The first two changes were addressed earlier. The estimate of the epistemic uncertainty in
581 the non-ergodic site term is described below.

582 **Epistemic Uncertainty in $\delta S2S$**

583 The standard deviation of the epistemic uncertainty of the non-ergodic term for the Cascadia-
584 specific non-ergodic model for the Puget Lowland shown in Figure 36 is computed using the
585 VCM results. For sites in the simulation region, the statistical estimate of the epistemic
586 uncertainty of the non-ergodic site term are all small (< 0.1 natural log units) due to the
587 large number of realizations (30) and the dense sampling of the site locations for the available
588 simulation data. The epistemic uncertainty from the VCM model is denoted σ_{mu-SIM} . For
589 sites outside the region with simulations, the epistemic uncertainty increases and reaches a
590 value close to ϕ_{S2S} . That is, for sites that are much farther from the simulation region than
591 the correlation length, the site term is assumed to be a random sample from the distribution
592 of $\delta S2S$ for the region with the simulations. The $\delta S2S$ are normally distributed with mean
593 zero and standard deviation ϕ_{S2S} .

594 With 30 realizations, the standard error of the non-ergodic site term in the simulation
595 region is small; however, this small uncertainty does not capture the uncertainty in the 3-
596 D velocity model. Ideally, we would have simulations for multiple 3-D velocity models to
597 capture the effects of the uncertainty in the 3-D velocity model, but there are high compu-
598 tational costs involved in these 3-D simulations. Currently, the M9 simulations are available
599 for only a single 3-D velocity model. This is also a limitation for other 3-D simulations for
600 large regions such as the 3-D simulations for the San Francisco Bay Area by Rodgers et al.
601 (2020).

602 Without results from a suite of 3-D velocity models, we assumed that the epistemic
603 uncertainty due to the 3-D velocity model is one half of the ϕ_{S2S} at all locations for $T >$
604 2 sec. In terms of variance, this is equivalent to assuming that the ground motions for the

alternative 3-D velocity models would have 25% of the variance of the $\delta S2S$ from the single velocity model. For shorter periods, we assume that the uncertainty due to the 3-D velocity model increases to 100% of the variance of $\delta S2S$ at $T=0.5$ sec (i.e., the 3-D velocity model does not provide any constraints on the site terms for $T \leq 0.5$ sec). With these assumptions, the variance of the epistemic uncertainty due to the 3-D velocity model is given by:

$$\sigma_{3D-VS}^2(T) = \phi_{S2S}^2(T)(1 - 0.75Taper_2(T)) \quad (41)$$

in which $Taper_2(T)$ is a linear taper (on a log period axis) from 0.5 sec to 2 sec:

$$Taper_2(T) = \begin{cases} 1 & \text{for } T > 2 \\ \left(1 + \frac{\ln(T/2)}{\ln(0.5/2)}\right) & \text{for } 0.5 < T \leq 2 \\ 0 & \text{for } T < 0.5 \end{cases} \quad (42)$$

This is an assumed model for the uncertainty in the ground motion due to the 3-D velocity model uncertainty; it has no quantitative basis from 3-D simulations with uncertain 3-D velocity structures. This assumed model should be reviewed and updated as additional 3-D simulations with alternative 3-D profiles become available.

For non-ergodic GMMs that include path-specific effects separate from the site effects, a large number of branches of the logic tree is needed to capture the uncertainty in the non-ergodic terms for all different combinations of ray paths: for example, Sung et al. (2021) and Lavrentiadis et al. (2021) use 100 branches to capture the uncertainty of the non-ergodic terms for combinations of all the different path effects for different source locations. For the partially non-ergodic GMM with only non-ergodic site terms, fewer branches are needed for the uncertainty in the non-ergodic because the non-ergodic terms apply to all source locations.

For the partially non-ergodic GMM, we discretize the normal distribution for the non-ergodic site terms that was developed as part of the VCM. In PSHA studies, it is common to model a normal distribution by a three-branch logic tree with site terms given by the median

shifted by -1.65, 0, and +1.65 times the epistemic uncertainty with weights of 0.185, 0.63, and 0.185. These shifts and weights recover the mean and variance of the underlying normal distribution. In the following example hazard analysis, we use this three-point discretization of the normal distribution for the logic tree branches for the epistemic uncertainty in the non-ergodic terms.

Example Sites

We selected four sites, shown in Figure 37, to use as examples for the comparison of the ergodic and non-ergodic hazard results. The coordinates and $Z_{2.5}$ values for the sites are listed in Table 8. SITE 1 is near the middle of the Seattle basin and has a large basin depth (6713 m). The locations of SITE 2 and SITE 3 are also in the Seattle basin, one near the east-northern area (SITE 2, basin depth = 2120 m), another close to the west-northern region (SITE 3, basin depth = 3295 m). The median non-ergodic site term for SITE 2 is a positive value, whereas the site term for SITE 3 is a negative value. The last site, SITE 4, is located outside the region for the simulations (basin depth = 1721 m). The median non-ergodic site term for SITE 4 is near zero. The location of the four sites on the map of the total basin term is shown in Figure 38. The total basin effects are largest at SITE 1 and SITE 2.

The distributions of the epistemic uncertainty for the four sites are shown in Figure 39. The distributions for the regions within the simulation region are much narrower than for SITE 4 which is outside the simulation region, reflecting the lack of simulation data to constrain the non-ergodic site term at SITE4.

Hazard Results

We calculated non-ergodic hazard for PSA at T=3 sec for the four example sites using the original AG20 GMM, the $AG20_{SIM-erg}$ GMM (with the modified basin scaling), the non-ergodic GMM with the regional σ_{SS} , and the non-ergodic GMM with a site-specific $\sigma_{SS,s}$.

For the non-ergodic GMM with the regional σ_{SS} , the 5th, 50th, and 95th fractiles for the epistemic uncertainty in the non-ergodic site term are also shown. These correspond to the three branches of the logic tree for the epistemic uncertainty in the non-ergodic site terms. For reference, the hazard for the BCHydro2016 GMM is also shown. The non-ergodic site terms and the aleatory standard deviations for each case are listed in Table 9.

For this calculation, we only included the hazard from the interface sources. The source model is taken from the BCHydro (2012) source model for Cascadia interface. Because the hazard is only for the interface sources, the hazard at 1E-4 to 2E-4 are likely to represent the ground-motion level for the 2500-yr return period from all source types (interface, slab, and crustal) based on the interface sources contributing 20-50% to the hazard. The changes to the 1E-4 ground motion shown in these plots will be larger than for the total hazard for all sources because only the interface GMM was modified; the hazard from the other sources (crustal and slab) will be unchanged.

The hazard curves are shown in Figure 40. Overall, non-ergodic hazard curves are steeper than others for all sites because the aleatory variability is reduced for the non-ergodic GMM. The non-ergodic hazard curves are also shifted to the left or right based on the positive or negative values of the site-specific non-ergodic term. The mean non-ergodic GMM includes the effect of the epistemic uncertainty in the non-ergodic terms. As this uncertainty increases, the slope of the mean hazard curve becomes flatter.

SITE 1 ($Z_{2.5} = 6713$ m with a negative (-0.09) non-ergodic term) shows a reduction in the mean non-ergodic hazard levels at low exceedance rates with a steeper slope because of the reduced aleatory variability. The stronger basin-depth scaling in $f_{basin3d}$ model compared to the original f_{basin} leads to an increase in the ground motion of about 10% for this basin depth. At an annual frequency of exceedance of 1×10^{-4} , the non-ergodic GMM leads to about a factor of 2 reduction in the hazard (y-axis). The hazard curve based on the BCHydro2016 model is much lower due to the lack of basin depth scaling in this GMM. For this site, the site-specific $\sigma_{SS,s}$ is similar to the regional σ_{SS} , so the two mean non-ergodic hazard curves

are similar.

For SITE 2 ($Z_{2.5} = 2120$ m with a positive (0.49) non-ergodic term), the hazard is similar for the AG20 and AG20 with 3-D basin terms because the $f_{basin3d}$ model is normalized to be equal to the f_{basin} model in the AG20 GMM for $Z_{2.5} = 2000$ m. The two non-ergodic mean hazard curves are much higher than the ergodic hazard curves due to the large positive non-ergodic site term. In this case, the site-specific $\sigma_{SS,s}$ is larger than the regional σ_{SS} for this basin edge site, so the site-specific single-station sigma leads to an increased hazard. The lower range of the epistemic fractiles for the non-ergodic hazard encompass the ergodic hazard curve.

For SITE 3 ($Z_{2.5} = 3295$ m with a negative (-0.71) non-ergodic term), the hazard is slightly larger for AG20 with 3-D basin terms than for the AG20 GMM because of the stronger basin-depth scaling in the 3-D simulations. The two non-ergodic mean hazard curves are much lower than the ergodic hazard curves due to the large negative non-ergodic site term. In this case, the site-specific $\sigma_{SS,s}$ is smaller than the regional σ_{SS} , so the site-specific sigma leads to a slightly steeper hazard curve. The upper range of the epistemic fractiles for the non-ergodic hazard is well below the ergodic hazard curves, indicating greater confidence in the change in the hazard for this site compared to the other sites.

For SITE 4 ($Z_{2.5} = 1721$ m with a near zero (-0.01) non-ergodic term), the hazard is lower for AG20 with 3-D basin terms than for the AG20 GMM because of the stronger basin-depth scaling is extended to $Z_{2.5}$ values less than 2000 m. The two non-ergodic mean hazard curves are similar to the ergodic hazard curves due to the large epistemic uncertainty in the non-ergodic site term that offsets the reduction in the aleatory variability. In this case, the site-specific $\sigma_{SS,s}$ is much larger than the regional σ_{SS} , so the site-specific sigma leads to a flatter hazard curve. The range of the epistemic fractiles for the non-ergodic hazard captures the ergodic hazard curves.

Comparing change in the hazard curves for the four sites, the T=3 sec ground motion at the $1E-4$ hazard level using the partially non-ergodic GMM (purple and orange dashed

curves) are up to a factor of 1.4 higher or a 2.5 lower than using the $AG20_{SIM-erg}$ ergodic model (green dashed curves). These hazard curves are only for the interface sources. If the interface contributes to 50% of the $T=3$ sec hazard, then this range would correspond to about a factor 1.25 increase or decrease in the 2500-yr ground motions for $T=3$ sec for the total hazard.

BASIN FACTORS FOR INTRA-SLAB EVENTS

The modified basin scaling model, $f_{basin3d}$, is based on the 3-D simulations for interface events and reflects the 3-D path effects for large interface events; however, given the sparse data for intra-slab earthquakes in the AG20 data set, we consider if the 3-D simulations for interface events can be used to improve the basin model for slab events. There are two questions about the basin model for intra-slab events to consider:

- Should the modified basin-depth scaling developed for the 3-D interface simulations, $f_{basin3d}$, be used for intra-slab events?
- Do the non-ergodic site terms developed for the interface site/path effects provide improved estimates for intra-slab events?

Basin-Depth Scaling for Intra-slab Events

The basin scaling from the 3-D simulations shows that the basin scaling extends to small basin depths ($Z_{25.5} < 2000$ m for $V_{S30}=600$ m/s) as compared to the AG20 basin scaling which limits the basin scaling to deep basin sites. To evaluate if the basin effects for intra-slab events should be extend to the $Z_{2.5} < 2000$ m, similar to the 3-D simulations, we first computed the residuals for the Cascadia intra-slab data using the modified basin scaling model and compared them to the residuals based on the AG20 basin scaling.

The mean residuals for these two basin models are shown in Figure 41 for all sites and for $Z_{2.5} > 2000$ m. For all sites, the AS20 basin scaling has a smaller mean residual than the

basin scaling based on the 3-D simulations. This is expected because the scaling in the AG20 model was developed by fitting these data. For the deeper basin sites, the mean residuals are similar between 2 and 6 seconds. The main difference is for the basin scaling for $Z_{2.5} < 2000$ m.

For the deeper basin sites, the slopes of the basin scaling between the unmodified f_{basin} model from AG20 and the modified $f_{basin3d}$ model are similar. The applicability of the modified basin scaling based on the 3-D simulations to basin effects for intra-slab earthquakes comes down to the expected behavior for smaller $Z_{2.5}$ values (less than 2000 m). Should the basin scaling for slab events be extended to shallow basin depths or is the scaling for shallow sites already captured by the V_{S30} scaling in the GMM due to the correlation between $Z_{2.5}$ and V_{S30} in the empirical data? We think that it is more likely that the basin effects from intra-slab earthquakes will extend to shallow basins in a similar manner as seen in the simulations for interface events than having very different scaling for shallow basin sites for intra-slab events. Therefore, we recommend using the $f_{basin3d}$ basin model for intra-slab events. If 3D simulations for slab events are available, then they should be used to check if the basin scaling for intra-slab events extends to shallow basin depths.

Applicability of Non-Ergodic terms to Intra-slab events

The non-ergodic terms for the AG20 interface model developed in this study are only based on the 3-D interface ground motions from M9 earthquakes. The non-ergodic terms are a combination of the site effect and the path effect for large interface earthquakes. We cannot separate the site and path effects, therefore, the non-ergodic terms are given for a specific site.

Even though the non-ergodic terms are for paths from interface events, we checked if the non-ergodic terms for interface events are correlated with the empirical site terms for intra-slab events. The residuals for intra-slab events from the unmodified AG20 GMM (without the ADJ_{CAS} term) are used. The site residuals, $\delta S2S$, from AG20 are adjusted to be relative

755 to the modified basin model:

$$\delta S2S' = \delta S2S_{AG20} - f_{basin}(Z_{2.5}) + f_{basin3d}(Z_{2.5}) \quad (43)$$

756 The modified site terms, $\delta S2S'$, from intra-slab events for periods of 3, 5, and 10 sec
 757 are plotted against the non-ergodic site/path terms for the interface simulations in Figure
 758 42. The period dependence of the correlation is shown in Figure 43. At periods less than
 759 3 sec, there is almost no correlation, indicating the 3-D path effects from the interface events
 760 are significantly different from the path effects for intra-slab events. At very long periods
 761 ($T=5-10$ sec), the correlation is larger, indicating that the site effects part of the combined
 762 path/site effects for interface events is significant for periods greater than 5 sec.

763 We recommend using reduced non-ergodic terms from the interface for application to
 764 intra-slab events. The reduction factor is the smoothed correlation coefficient, ρ , as shown
 765 in equation (44):

$$\delta S2S_{nonerg-slab}(T, t_s) = \delta S2S_{nonerg}(T, t_s)\rho(T) \quad (44)$$

766 in which

$$\rho(T) = \begin{cases} 0 & \text{for } T < 3 \\ 0.5 \left(\frac{\ln(T/3)}{\ln(5/3)} \right) & \text{for } 3 < T < 5 \\ 0.5 & \text{for } T > 5 \end{cases} \quad (45)$$

767 For intra-slab events, the epistemic uncertainty in the non-ergodic site terms will be
 768 larger than for the interface events because the correlation between the intra-slab residuals
 769 and the non-ergodic site terms from the interface events is less than unity. The epistemic
 770 uncertainty in the non-ergodic site term for intra-slab events due to the uncertainty in the
 771 3-D velocity structure is given by:

$$\sigma_{3D-VS-slab}^2(T) = \phi_{S2S}^2(T)(1 - 0.75Taper_2(T)\rho^2(T)) \quad (46)$$

The epistemic uncertainty for the intra-slab non-ergodic site terms can be written in terms of the epistemic uncertainty for the interface events as shown in equation 47. As the correlation, $\rho(T)$, goes to zero, the epistemic uncertainty goes to the full ϕ_{S2S} value.

$$\sigma_{3D-VS-slab}^2(T) = \sigma_{3D-VS}^2(T) + 0.75(1 - \rho^2(T)) Taper_2(T) \phi_{S2S}^2 \quad (47)$$

APPLICATION GUIDELINES

For implementing non-ergodic GMMs into PSHA, the non-ergodic GMM needs to include three terms: the adjustment to the median, the reduced aleatory variability, and the epistemic uncertainty in the adjustment term. The non-ergodic AG20 GMM developed in this study provides includes these three terms for large-magnitude interface events for Cascadia. While the non-ergodic terms were developed only for M9 earthquakes, they are assumed to apply to all interface scenarios (all magnitudes and locations). Given that the contribution from the interface sources to the seismic hazard in the Seattle region is dominated by the larger events that rupture the part of the interface close to Seattle, this is considered a reasonable assumption for use in PSHA for the Seattle region.

A concern with non-ergodic GMMs is that they are only for limited regions, but the seismic hazard needs to be computed for larger regions that are not covered by the 3-D simulations. The framework of the non-ergodic GMM can still be used outside the regions with simulations: for sites outside the simulation region, the reduced aleatory variability (σ_{SS}) is still used with a zero median adjustment but with large epistemic uncertainty in the non-ergodic site terms. If the epistemic uncertainty in the median adjustment is set at ϕ_{S2S} , then the mean hazard using the non-ergodic GMM will be equal to the mean hazard using the ergodic GMM with no adjustment and with the ergodic σ , but the uncertainty fractiles will be wider, reflecting the lack of data to constrain the site terms. This allows for a consistent approach to be used for the entire region even if the 3-D simulations are limited

to a few subregions.

The epistemic uncertainty in GMMs is more than the uncertainty in the non-ergodic terms about a reference GMM. The uncertainty in the base GMM should also be included. The integration of the current set of M9 simulations for Cascadia into a non-ergodic version of the AG20 GMM allows provides one GMM that can be used in PSHA calculations. The modification to the AG20 model to make it a non-ergodic model does not change the average level of the ground motion; it only changes the spatial distribution of the site terms and the aleatory variability. Other GMMs, with different magnitude scaling and overall ground-motion levels can also be modified in a similar manner to capture the range of GMMs. Alternatively, the scaled-backbone approach can be used by scaling the non-ergodic AG20 GMM to capture the epistemic range from alternative GMMs.

Our recommendation for implementation of the non-ergodic AG20 GMM for Cascadia is summarized below. As mentioned earlier, this approach can be applied to all sites in the Cascadia region, not just to the sites in the subregion with the with the 3-D simulations.

Median

The median ground motion for interface events is given by adding the non-ergodic site term to the $AG20_{SIM-erg}(M, R, \dots)$ median (from equation 14):

$$\mu_{inter}(M, R, \dots, Lat_s, Long_s) = AG20_{SIM-erg}(M, R, \dots) + \delta S2S_{nonerg}(T, Lat_s, Long_s) \quad (48)$$

The median ground motion for intra-slab events is given by adding the non-ergodic site term for slabs (reduced interface site term) to the $AG20_{SIM-erg}(M, R, \dots)$ median:

$$\mu_{slab}(M, R, \dots, Lat_s, Long_s) = AG20_{SIM-erg}(M, R, \dots) + \rho(T)\delta S2S_{nonerg}(T, Lat_s, Long_s) \quad (49)$$

Aleatory variability

The single-station sigma is given by either $\sigma_{SS}(T)$ or $\sigma_{SS,s}(T, Lat_s, Long_s)$. The τ and ϕ_{SS} are from the global models given by AG20 GMM (equations 33 and 37). The global ϕ_{SS} is recommended for use because Abrahamson and Gulerce (2020) showed that ϕ_{SS} is not region dependent. If the site-specific single-station sigma is used, then the $\sigma_{SS,s}(T, Lat_s, Long_s)$ values from the simulations are scaled by the ratio of $\phi_{SS-AG20}/\phi_{SS-SIM}$ to be consistent with the global ϕ_{SS} .

Epistemic uncertainty

For interface events, the epistemic uncertainty in the non-ergodic term is given by

$$\sigma_{nonerg-inter}(T, Lat_s, Long_s) = MAX(\sigma_{\mu-SIM}(T, Lat_s, Long_s), \sigma_{3D-Vs}(T)) \quad (50)$$

For intra-slab events, the epistemic uncertainty in the non-ergodic term is given by:

$$\sigma_{nonerg-slab}(T, Lat_s, Long_s) = MAX(\sigma_{\mu-SIM}(T, Lat_s, Long_s), \sigma_{3D-Vs-slab}(T)) \quad (51)$$

The ϕ_{S2S} model used to compute the σ_{3D-Vs} and $\sigma_{3D-Vs-slab}$ terms is the Cascadia-specific model given in Table 6.

This does not capture the epistemic uncertainty in the reference ergodic GMM (the $AG20_{SIM-erg}$ GMM in this case). This uncertainty can be captured by using alternative reference GMMs or using the scaled-backbone approach with the $AG20_{SIM-erg}$ GMM.

Coefficients

The values of the spatially varying terms ($\delta S2S(T, Lat_s, Long_s)$, $\sigma_{mu-SIM}(T, Lat_s, Long_s)$, and $\sigma_{SS,s}(T, Lat_s, Long_s)$) are given in the file "AG20-nonerg-coeff.xlsx". For sites outside

the region included in the file, use the following:

$$\delta S2S_{nonerg}(T, Lat_s, Long_s) = 0,$$

$$\sigma_{SS,s}(T, R_{RUP}, Lat_s, Long_s) = \sigma_{SS}(T, R_{RUP})$$

$$\sigma_{mu-SIM}(T, Lat_s, Long_s) = \phi_{S2S}(T)$$

830 This approach leads to a mean hazard that is the same as the hazard using the ergodic
 831 AG20 GMM with σ based on the Cascadia-specific ϕ_{S2S} rather than the AG20 ϕ_{S2S} , but the
 832 epistemic fractiles are broader which reflect the lack of data to constrain the non-ergodic
 833 basin terms.

834 Hazard Calculation

835 Implementing this model into PSHA calculations requires modifying the aleatory standard
 836 deviation to be σ_{SS} , including site-specific adjustments to the median based on latitude
 837 and longitude of the site. An additional logic tree node with the alternative values of the
 838 site-specific adjustment is needed.

839 The hazard is computed using the conditional probability of exceeding the ground motion
 840 given by:

$$P(Y > z|M, R...) = 1 - \Phi \left(\frac{\ln(z) - [AG20_{SIM-erg}(M, R, ...) + \delta S2S(Lat_s, Lon_s)]}{\sigma_{SS}(M, R, Lat_s, Lon_s)} \right) \quad (52)$$

841 The non-ergodic adjustment, $\delta S2S(Lat_s, Lon_s)$, is given by:

$$\delta S2S(Lat_s, Lon_s) = \delta \hat{S}2S(Lat_s, Lon_s) + \delta S2S_{epi}(Lat_s, Lon_s) \quad (53)$$

The $\delta\hat{S}^2S(Lat_s, Lon_s)$ is the median estimate of the non-ergodic term. The digital values for $\delta\hat{S}^2S(Lat_s, Lon_s)$ are in sheet "01_nonerg_mean_vcm" of file "Cascadia_NonergTerm&SigmaSS_Oct29.xls".

The $\delta S^2S_{epi}(Lat_s, Lon_s)$ is the uncertainty in the non-ergodic term from the logic tree sampling of the normal distribution with mean zero and standard deviation $\sigma_{nonerg-inter}$ for interface events. The standard deviation of the epistemic uncertainty of the non-ergodic term is given in sheet "03_nonerg_epis_(0.5_S2S+vcm)" of file "Cascadia_NonergTerm&SigmaSS_Oct29.xlsx". This value includes the estimate of the uncertainty due to the uncertainty in the 3-D velocity structure. The uncertainty due only to the VCM is given in sheet "02_nonerg_epis_vcm" for reference, but this value should not be used by itself. If there is an improved estimate of the uncertainty due to the 3-D velocity model, then the values in sheet "02_nonerg_epis_vcm" can be combined with the updated 3-D velocity model uncertainty.

The single-station sigma term, $\sigma_{SS}(M, R, Lat_s, Lon_s)$, can be an average value for the region or a site-specific value. The period dependence of the average σ_{SS} is given in the "AG20 Non-ergodic" column of Table 7. For periods less than 1 sec, the sigma from the $AG20_{SIM-erg}$ column of Table 7 can be used. The site-specific values of $\sigma_{SS}(M, R, Lat_s, Lon_s)$ are given in sheet "04_SigmaSS,S" of the file "Cascadia_NonergTerm&SigmaSS_Oct29.xlsx".

CONCLUSIONS

As sets of 3-D simulations become more widely available, there will be a need to incorporate the information from the 3-D simulations into ground-motion models used in seismic hazard calculations. With limited sets of scenarios used in the 3-D simulations expected for the most regions, the partially non-ergodic GMM approach for the site term is the most practical approach for including results from 3-D simulations into GMMs at this time. As additional 3-D simulations for suites of scenarios with different paths become available, non-ergodic models that include site/source-specific path effects, as well as site-specific site effects, can be developed.

The 3-D simulations can be incorporated into both the median and aleatory standard deviation of GMMs to capture non-ergodic basin effects. The reference-station approach for estimating basin effects can be inconsistent with the V_{S30} scaling in the GMM that may implicitly include $Z_{2.5}$ scaling, which can lead to double counting of basin effects in the modified GMM. The approach of using a normalized basin depth, in which the normalization varies with the V_{S30} , avoids the potential for double counting basin effects already included in the V_{S30} scaling.

When using partially non-ergodic GMMs, it is important to include the epistemic uncertainty in the non-ergodic terms as part of the logic tree. A key limitation of the current model is that the epistemic uncertainty in the non-ergodic basin terms due to uncertainty in the 3-D velocity model is based on assumed values without the benefit of simulation results for a suite of alternative of 3-D velocity models to constrain the estimates. This is a topic that needs to be addressed if 3-D simulation results are to be used in seismic hazard practice and not just in illustrative examples.

Data and Resources

The response spectra from the M9 simulations are available from the DesignSafe website at [https://doi.org/ 10.17603/DS2WM3W](https://doi.org/10.17603/DS2WM3W) (last accessed April 2021).

Acknowledgments

This study was funded by the USGS NEHRP program, award number G20AP00054. We thank Ellen Rathje for help with downloading the full set of response spectra for the M9 simulations.

References

- Abrahamson, N., N. Gregor, and K. Addo (2016). Bc hydro ground motion prediction equations for subduction earthquakes. *Earthquake Spectra* 32(1), 23–44.
- Abrahamson, N. and Z. Gulerce (2020). Regionalized ground-motion models for subduction earthquakes based on the nga-sub database. Technical report, PEER Rept. 2020/25.
- Al Atik, L., N. A. Abrahamson, J. J. Bommer, F. Scherbaum, F. Cotton, and N. M. Kuehn (2010). The Variability of Ground-Motion Prediction Models and Its Components. *Seismological Research Letters* 81(5), 794–801.
- Frankel, A., E. Wirth, N. Marafi, J. Vidale, and W. Stephenson (2018). Broadband synthetic seismograms for magnitude 9 earthquakes on the cascadia megathrust based on 3d simulations and stochastic synthetics, part 1: Methodology and overall results. *Bulletin of the Seismological Society of America* 108(5A), 2347–2369.
- Kuehn, N., Y. Bozorgnia, K. Campbell, and N. Gregor (2020). Partially nonergodic ground-motion model for subduction regions using nga-subduction database. *PEER Rept.* 2020/04.
- Landwehr, N., N. M. Kuehn, T. Scheffer, and N. Abrahamson (2016, 12). A Nonergodic Ground-Motion Model for California with Spatially Varying Coefficients. *Bulletin of the Seismological Society of America* 106(6), 2574–2583.
- Lavrentiadis, G., N. Abrahamson, and N. Kuehn (2021). A non-ergodic effective amplitude ground-motion model for california. *Bulletin of Earthquake Engineering*.
- Lin, P.-S., B. Chiou, N. Abrahamson, M. Walling, C.-T. Lee, and C.-T. Cheng (2011). Repeatable source, site, and path effects on the standard deviation for empirical ground-motion prediction models. *Bulletin of the Seismological Society of America* 101(5), 2281–2295.

912 Lindgren, F., H. Rue, et al. (2015). Bayesian spatial modelling with r-inla. *Journal of*
913 *Statistical Software* 63(19), 1–25.

914 McCrory, P. A., J. L. Blair, F. Waldhauser, and D. H. Oppenheimer (2012). Juan de fuca
915 slab geometry and its relation to wadati-benioff zone seismicity. *Journal of Geophysical*
916 *Research: Solid Earth* 117(B9).

917 Meng, X. and C. Goulet (2021). Lessons learned from applying varying coefficient model to
918 controlled simulation datasets. *Bulletin of Earthquake Engineering*.

919 Moschetti, M., M. Ritzwoller, F.-C. Lin, and Y. Yang (2010). Crustal shear wave velocity
920 structure of the western united states inferred from ambient seismic noise and earthquake
921 data. *Journal of Geophysical Research: Solid Earth* 115(B10).

922 Parker, G., J. Stewart, D. Boore, G. Atkinson, and B. Hassani (2020). Nga-subduction
923 global ground-motion models with regional adjustment factors. *PEER Report* 3.

924 Rodgers, A. J., A. Pitarka, R. Pankajakshan, B. Sjögreen, and N. A. Petersson (2020). Mw7
925 earthquakes on the hayward fault, northern california resolving frequencies 0–10 hz and
926 including site-response corrections. *Bulletin of the seismological society of America* 110(6),
927 2862–2881.

928 Rue, H., S. Martino, and N. Chopin (2009). Approximate bayesian inference for latent
929 gaussian models by using integrated nested laplace approximations. *Journal of the royal*
930 *statistical society: Series b (statistical methodology)* 71(2), 319–392.

931 Rue, H., A. Riebler, S. H. Sørbye, J. B. Illian, D. P. Simpson, and F. K. Lindgren (2017).
932 Bayesian computing with inla: a review. *Annual Review of Statistics and Its Application* 4,
933 395–421.

934 SDCI (2018). Implementation of march 22, 2018 usgs/sdci basin amplification workshop

935 results, director’s rule 20-2018. Technical report, Seattle Department of Construction and
936 Inspections.

937 Stan Development Team (2020). RStan: the R interface to Stan. R package version 2.21.2.

938 Stephenson, W. J., N. G. Reitman, and S. J. Angster (2017). P-and s-wave velocity models
939 incorporating the cascadia subduction zone for 3d earthquake ground motion simulations,
940 version 1.6—update for open-file report 2007–1348. Technical report, US Geological Sur-
941 vey.

942 Sung, C.-H., N. Abrahamson, N. Kuehn, P. Traversa, and I. Zentner (2021). A non-ergodic
943 ground-motion model of fourier amplitude spectra for france. *Bulletin of Earthquake*
944 *Engineering*.

945 Wirth, E. A., A. D. Frankel, N. Marafi, J. E. Vidale, and W. J. Stephenson (2018). Broad-
946 band synthetic seismograms for magnitude 9 earthquakes on the cascadia megathrust
947 based on 3d simulations and stochastic synthetics, part 2: Rupture parameters and vari-
948 ability. *Bulletin of the Seismological Society of America* 108(5A), 2370–2388.

949

950

951

952

953

Table 1: Epicenter coordinates and down-dip edge group for the 30 source realizations used in the 3-D simulations.

ID.no	Epicenter		Down-Dip Edge
	Latitude($^{\circ}$)	Longitude ($^{\circ}$)	
1	43.618	-125.417	middle
2	47.941	-125.531	easternmost
3	45.141	-125.255	middle
4	48.246	-124.667	middle
5	44.697	-124.913	middle
6	48.360	-126.060	westernmost
7	43.779	-124.903	westernmost
8	41.781	-124.123	easternmost
9	43.746	-123.982	easternmost
10	48.563	-126.278	middle
11	43.050	-124.988	middle
12	46.871	-124.544	middle
13	46.145	-124.942	middle
14	42.731	-124.627	westernmost
15	47.659	-125.327	middle
16	47.408	-124.786	middle
17	41.352	-124.728	middle
18	44.227	-124.656	middle
19	42.308	-124.858	middle
20	41.607	-124.376	middle
21	41.062	-124.722	middle
22	46.766	-125.439	westernmost
23	46.063	-124.761	westernmost
24	40.720	-124.682	westernmost
25	48.986	-126.373	easternmost
26	45.588	-125.341	easternmost
27	42.890	-124.901	easternmost
28	46.690	-124.392	easternmost
29	44.714	-124.590	easternmost
30	40.901	-124.420	easternmost

Table 2: Mean residuals by $Z_{2.5}$ bin shown in Figures 7 to 9 .

Z2.5' Bin	Z2.5 (m) Bin	Mean Residual (LN units)					
		T=0.2	T=1	T=2	T=3	T=5	T=10
≤ 0.073	≤ 100	0.347	-0.095	-0.380	-0.463	-0.599	-0.048
0.073 - 0.122	100 - 200	0.388	-0.059	-0.348	-0.464	-0.539	-0.049
0.122 - 0.171	200 - 300	0.380	-0.011	-0.354	-0.456	-0.547	0.024
0.171 - 0.269	300 - 500	0.337	0.024	-0.243	-0.375	-0.474	0.034
0.269 - 0.513	500 - 1000	0.331	0.159	-0.038	-0.218	-0.380	0.077
0.513 - 1.001	1000 - 2000	0.336	0.410	0.558	0.298	-0.063	0.246
1.001 - 1.489	2000 - 3000	0.359	0.635	1.032	0.845	0.336	0.471
1.489 - 2.466	3000 - 5000	0.382	0.683	1.266	1.130	0.663	0.751
2.466 - 4.907	5000 - 10000	0.420	0.679	1.363	1.350	0.962	1.044

Table 3: Coefficients for the basin-depth scaling).

T(sec)	b_1	b_2	Z'_x	C_{SIM}
1.0	0.264	-0.072	0.106	0.523
1.5	0.569	-0.310	0.132	0.843
2.0	0.725	-0.365	0.197	0.812
2.5	0.822	-0.407	0.250	0.731
3.0	0.872	-0.462	0.298	0.594
4.0	0.856	-0.572	0.347	0.333
5.0	0.790	-0.573	0.391	0.168
6.0	0.712	-0.485	0.418	0.137
7.5	0.653	-0.278	0.446	0.249
10. 0	0.575	-0.038	0.477	0.388

Table 4: Hyperparameters for the covariance function.

T(sec)	Θ (LN units)	ρ (km)
1.0	0.354	36.51
1.5	0.383	39.12
2.0	0.437	41.85
2.5	0.435	39.70
3.0	0.379	40.69
4.0	0.278	38.89
5.0	0.200	38.50
6.0	0.151	39.26
7.5	0.113	38.91
10.0	0.078	41.24

Table 5: Coefficients for the AG20 ϕ_1^2 model applicable to Cascadia.

Period (sec)	d_1	d_2
0.01	0.325	0.137
0.02	0.325	0.137
0.03	0.325	0.137
0.05	0.325	0.137
0.075	0.325	0.137
0.10	0.325	0.137
0.15	0.325	0.137
0.20	0.325	0.137
0.25	0.325	0.137
0.30	0.325	0.137
0.40	0.325	0.137
0.50	0.325	0.137
0.60	0.325	0.137
0.75	0.325	0.137
1.0	0.325	0.137
1.5	0.312	0.113
2.0	0.302	0.096
2.5	0.295	0.082
3.0	0.289	0.072
4.0	0.280	0.055
5.0	0.273	0.041
6.0	0.267	0.030
7.5	0.259	0.017
10.0	0.250	0.000

Table 6: Within-event and between site standard deviations based on the 3-D simulations and AG20 GMM for distances less than 150 km.

Period (sec)	From 3-D Simulations			AG20 GMM			Cascadia-Specific
	ϕ	ϕ_{SS}	ϕ_{S2S}	ϕ	ϕ_{SS}	ϕ_{S2S}	ϕ_{S2S}
0.000	0.219	0.201	0.087	0.570	0.41	0.396	0.396
0.020	0.218	0.200	0.086	0.570	0.41	0.396	0.396
0.030	0.215	0.198	0.082	0.570	0.41	0.396	0.396
0.050	0.205	0.195	0.064	0.570	0.41	0.396	0.396
0.075	0.251	0.210	0.138	0.570	0.41	0.396	0.396
0.100	0.185	0.184	0.024	0.570	0.41	0.396	0.396
0.150	0.180	0.178	0.026	0.570	0.41	0.396	0.396
0.200	0.183	0.176	0.049	0.570	0.41	0.396	0.396
0.250	0.190	0.176	0.070	0.570	0.41	0.396	0.396
0.300	0.198	0.178	0.086	0.570	0.41	0.396	0.396
0.400	0.199	0.178	0.089	0.570	0.41	0.396	0.396
0.500	0.200	0.179	0.088	0.570	0.41	0.396	0.396
0.600	0.204	0.184	0.089	0.570	0.41	0.396	0.396
0.750	0.230	0.209	0.097	0.570	0.41	0.396	0.396
1.000	0.375	0.299	0.226	0.570	0.41	0.396	0.396
1.500	0.523	0.351	0.387	0.559	0.41	0.379	0.387
2.000	0.539	0.335	0.423	0.550	0.41	0.366	0.423
2.500	0.533	0.316	0.429	0.543	0.41	0.356	0.429
3.000	0.520	0.307	0.420	0.538	0.41	0.348	0.420
4.000	0.471	0.286	0.375	0.529	0.41	0.335	0.375
5.000	0.340	0.266	0.294	0.522	0.41	0.324	0.309
6.000	0.337	0.241	0.296	0.517	0.41	0.314	0.305
7.500	0.310	0.236	0.200	0.509	0.41	0.301	0.251
10.000	0.291	0.253	0.145	0.500	0.41	0.286	0.216

Table 7: Total aleatory standard deviation for AG20, $AG30_{SIM-erg}$, AG20 non-ergodic, and the sigma model of Abrahamson and Gulerce (2020)

	ϕ from 3-D simulations, τ from AG20			Empirical ϕ and τ
T(sec)	$AG20_{SIM-erg}$ (w/modified basin)	AG20 Non-ergodic	AG20 (w/o modified basin)	AG20 GMM $R_{RUP} < 150$ km
0.2	0.504	-	0.503	0.739
0.25	0.507	-	0.505	0.739
0.3	0.510	-	0.508	0.739
0.4	0.510	-	0.514	0.739
0.5	0.511	-	0.519	0.739
0.6	0.513	-	0.523	0.739
0.75	0.523	-	0.535	0.739
1.0	0.601	0.554	0.639	0.739
1.5	0.703	0.582	0.851	0.730
2.0	0.716	0.574	0.887	0.723
2.5	0.711	0.563	0.883	0.718
3.0	0.701	0.558	0.860	0.714
4.0	0.666	0.547	0.790	0.708
5.0	0.615	0.537	0.702	0.703
6.0	0.579	0.526	0.640	0.698
7.5	0.563	0.525	0.608	0.693
10	0.553	0.532	0.585	0.686

Table 8: Locations of the four sites used in the example hazard calculations.

Site	Longitude	Latitude	$Z_{2.5}$ (m)	$f_{basin3d} - f_{basin}$ for $T=3$ sec
Site 1	-122.20	47.60	6713	0.1
Site 2	-122.70	47.80	2120	0.0
Site 3	-122.70	47.60	3295	0.05
Site 4	-123.60	48.10	1721	-0.15

Table 9: Ground-motion input parameters for $T=3$ sec for the example hazard calculations for the four sites. The standard deviation is simplified to be constant with distance less than 150 km)

		Non-Ergodic Term				
Site	GMM	Median	Epistemic Unc	σ	σ_{SS}	$\sigma_{SS,s}$
All sites	AG20	0.0	-	0.714	-	-
All sites	AG20 modified basin	0.0	-	0.701	-	-
Site 1	AG20 non-ergodic	-0.088	0.21	-	0.558	0.567
Site 2	AG20 non-ergodic	0.493	0.21	-	0.558	0.638
Site 3	AG20 non-ergodic	-0.715	0.21	-	0.558	0.530
Site 4	AG20 non-ergodic	-0.009	0.42	-	0.558	0.672

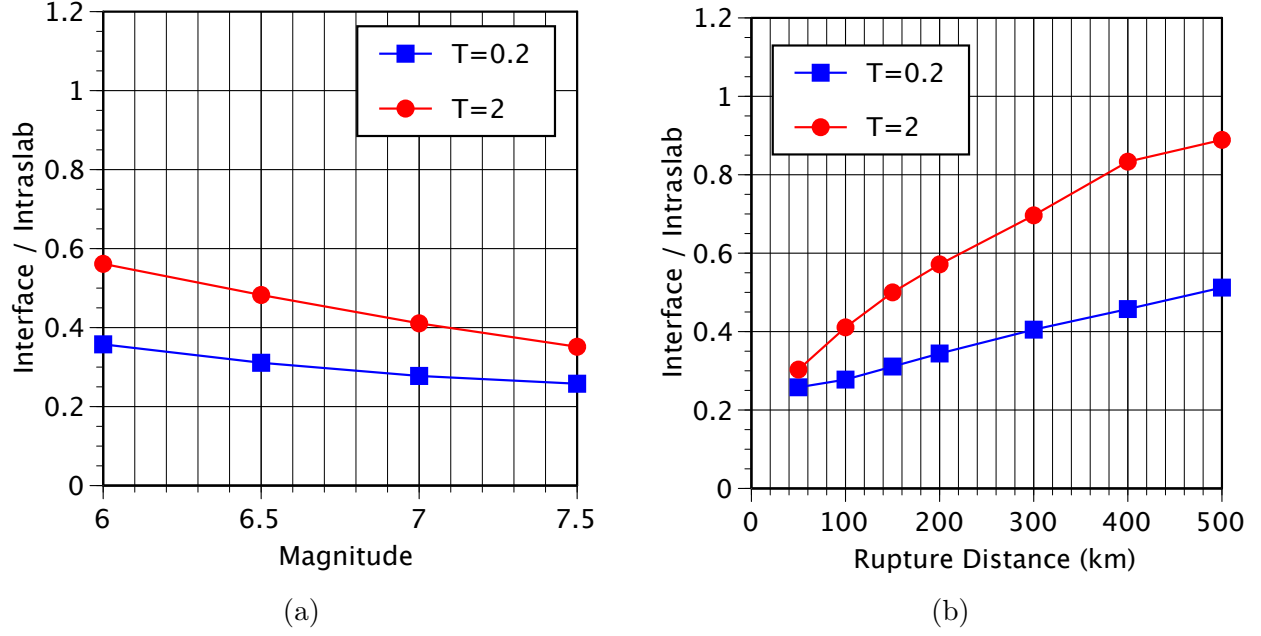


Figure 1: (a) Magnitude scaling of the interface/intra-slab ratio for $R_{RUP}=100$ km and $Z_{TOR}=50$ km. (b) Distance scaling of the interface/intra-slab ratio for $M=7$ and $Z_{TOR}=50$ km.

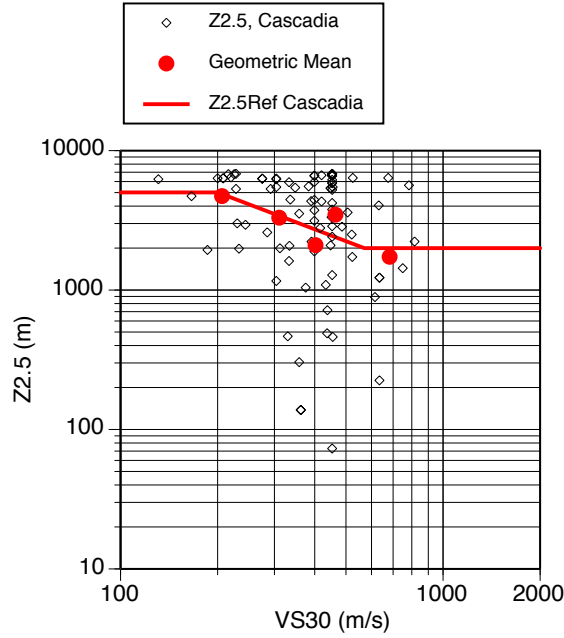


Figure 2: Correlation of $Z_{2.5}$ and V_{S30} for the Cascadia data used by AG20. The geometric mean of the $Z_{2.5}$ is shown by the circles and the model used for the Z_{ref} is shown by the line.

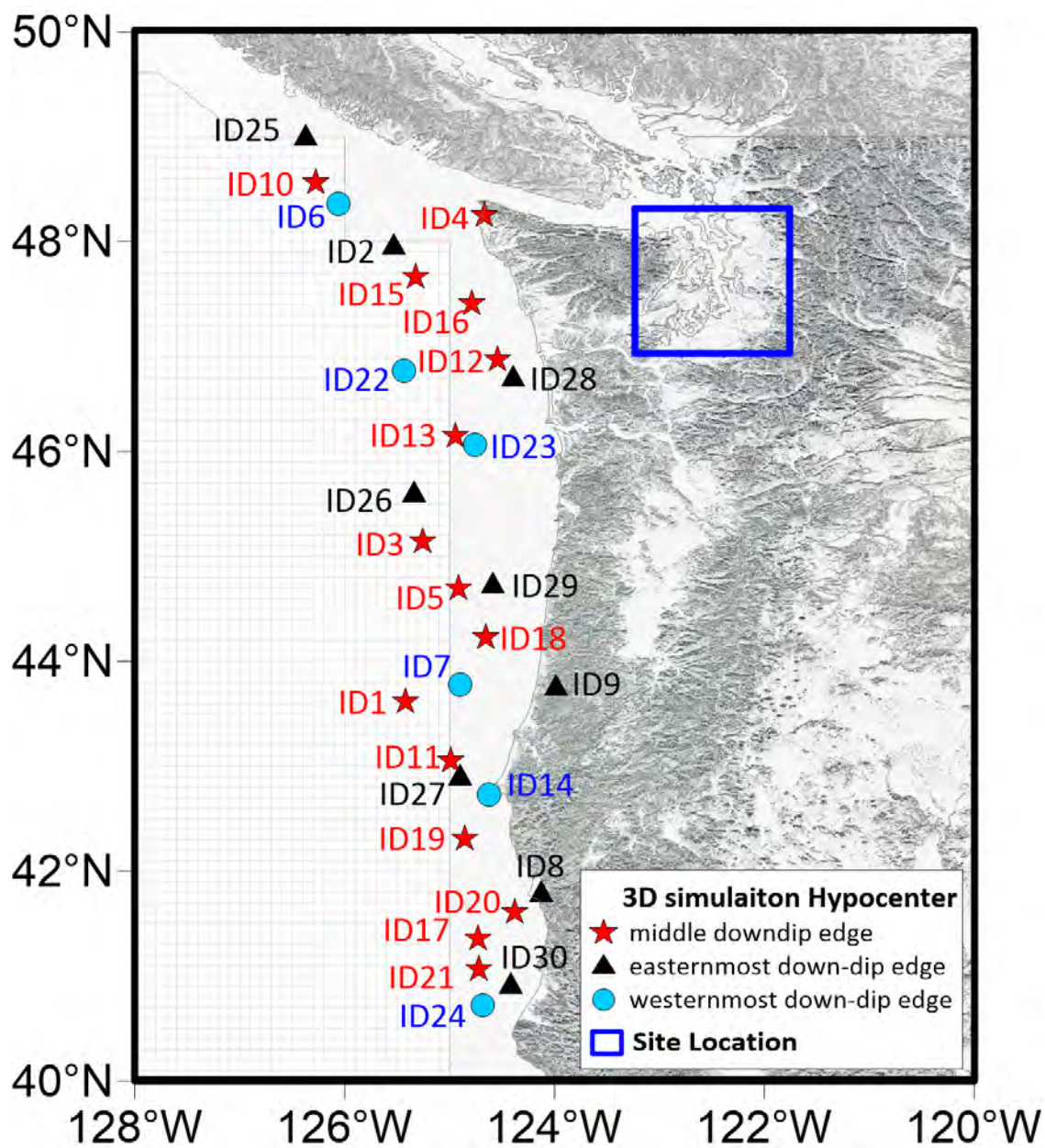


Figure 3: Distribution map of 30 simulations and site locations used in this study. Stars indicate three-dimensional hypocenters and a rectangle indicates site region.

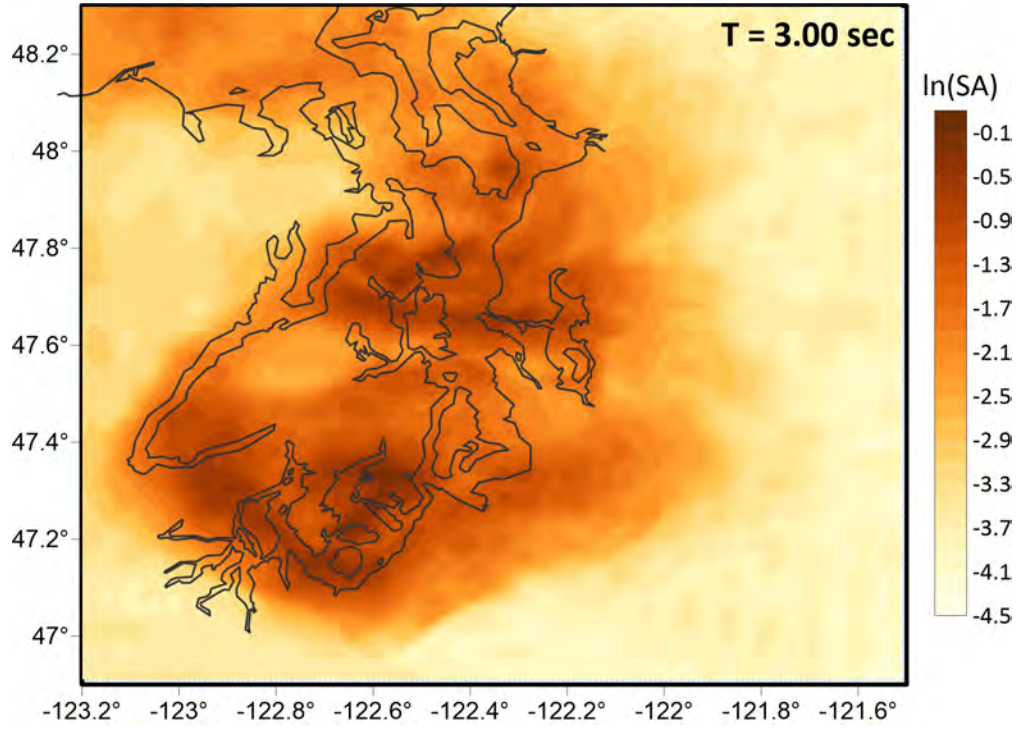


Figure 4: Map of the spectral acceleration at $T=3$ sec for the Puget Lowland from event ID1 (Frankel et al., 2018).

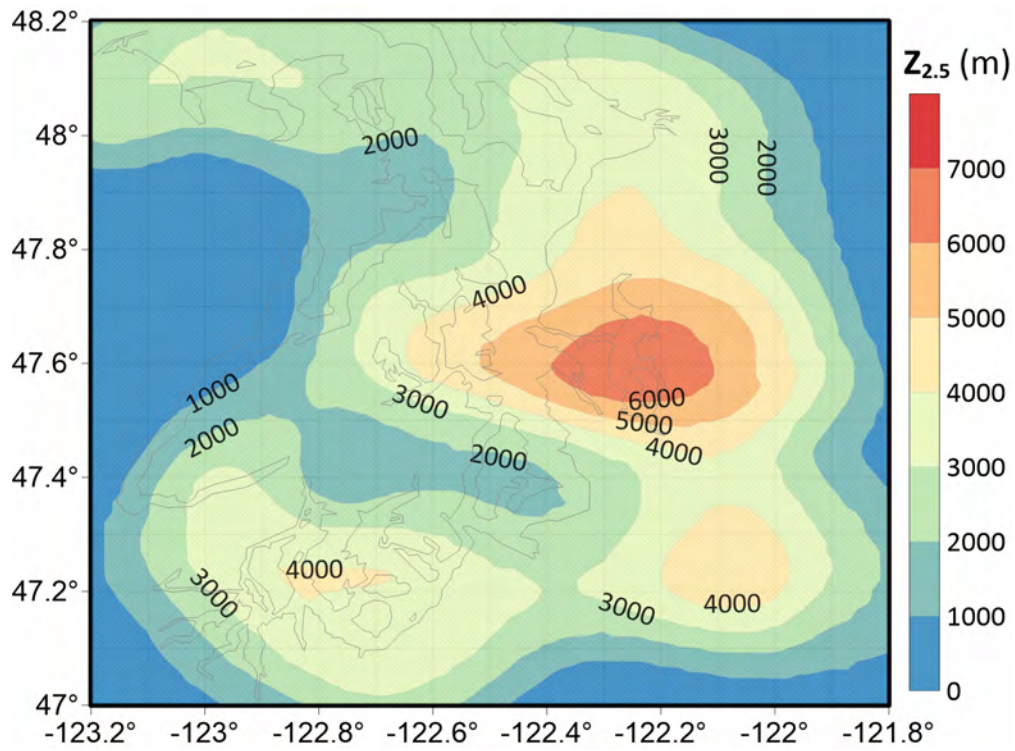


Figure 5: Distribution of basin-depth parameter ($Z_{2.5}$) for the study region.

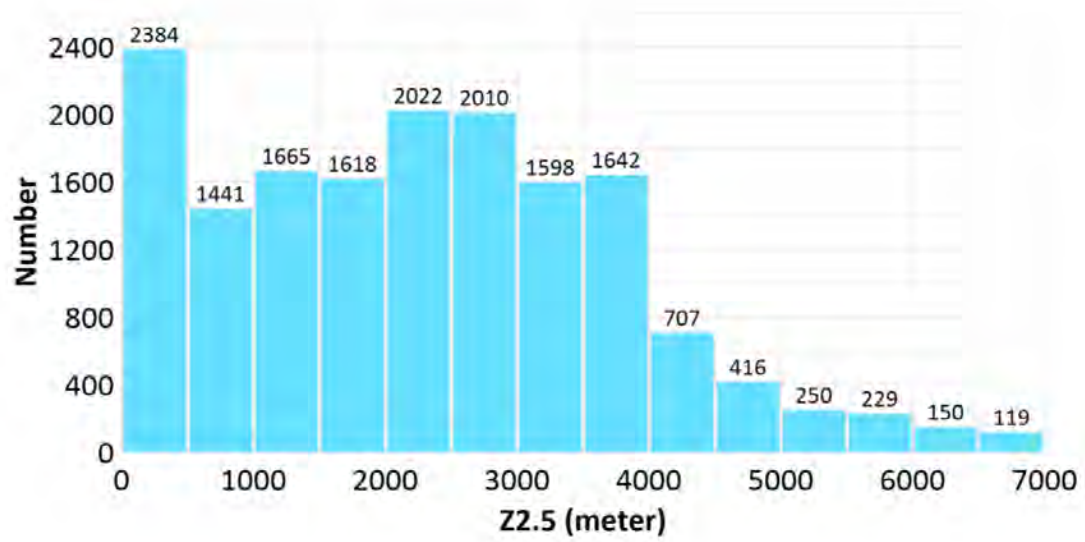


Figure 6: Distribution of the basin-depth parameter, $Z_{2.5}$, for the study region shown in Figure 5.

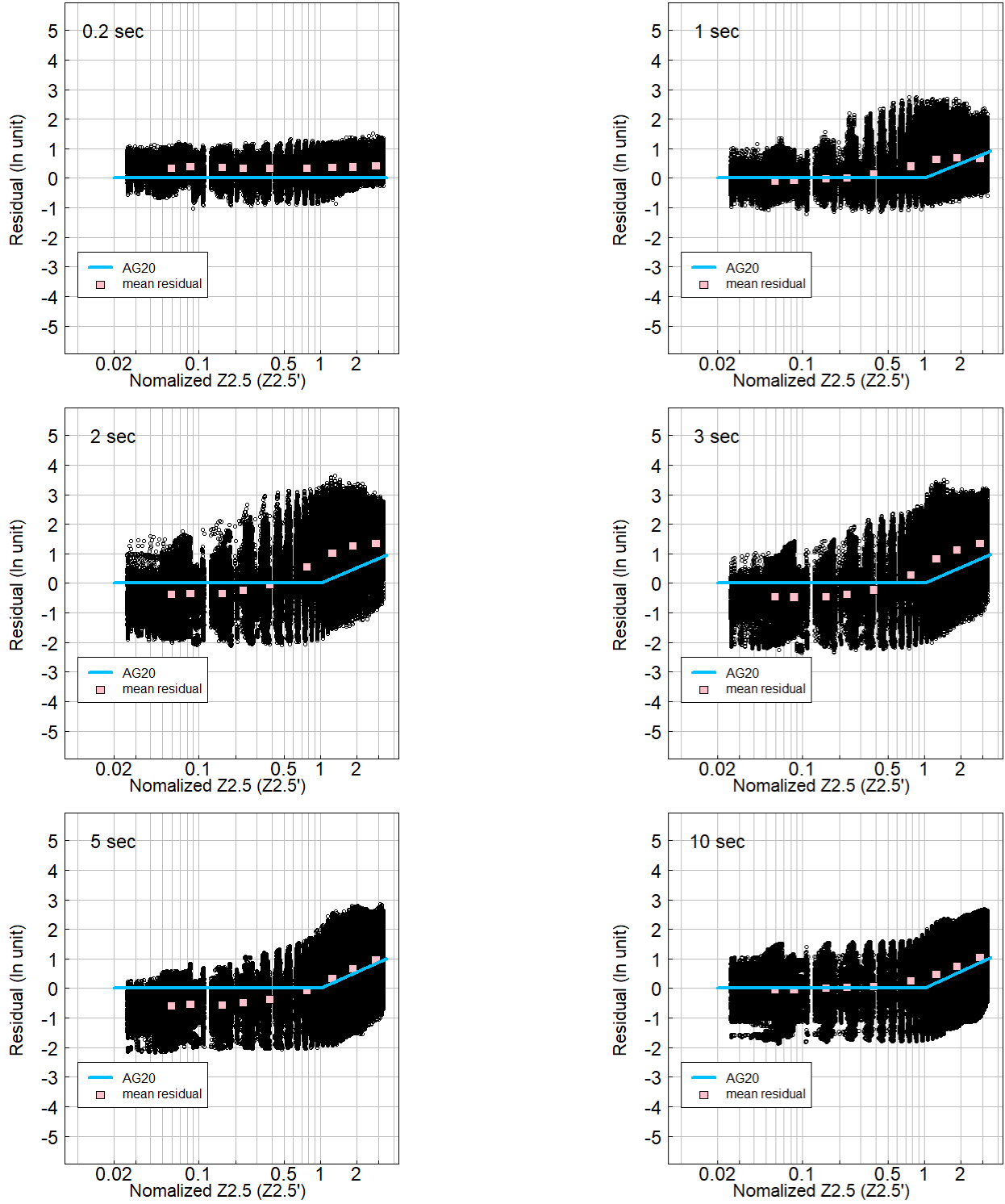


Figure 7: Total residuals of the M9 simulations relative to the AG20 model. The basin-depth scaling for the AG20 model is shown by the blue curve and the mean residuals from the M9 simulations for $Z'_{2.5}$ bins are shown by the pink squares.

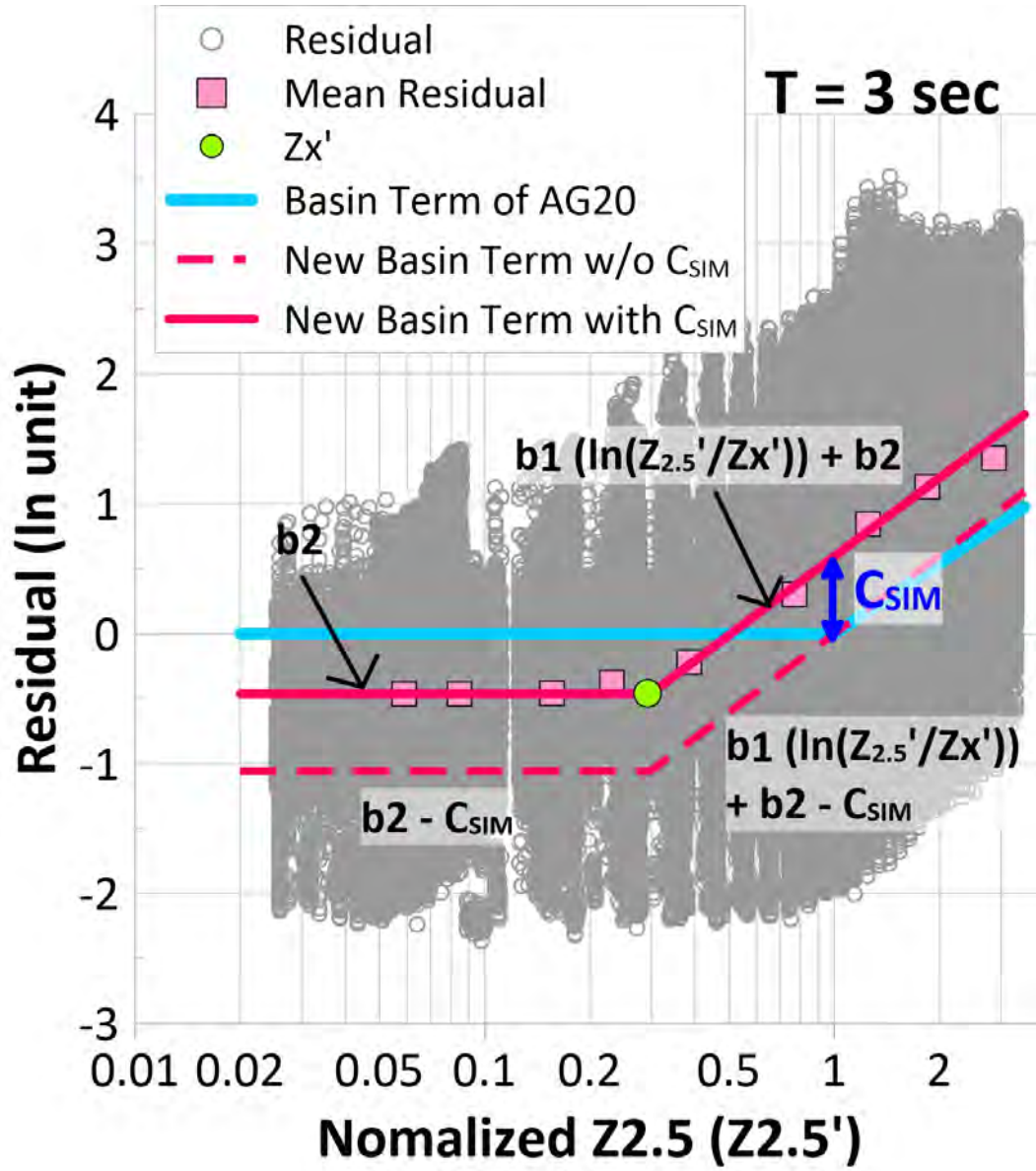


Figure 8: An example for new basin depth term at $T = 3$ sec.

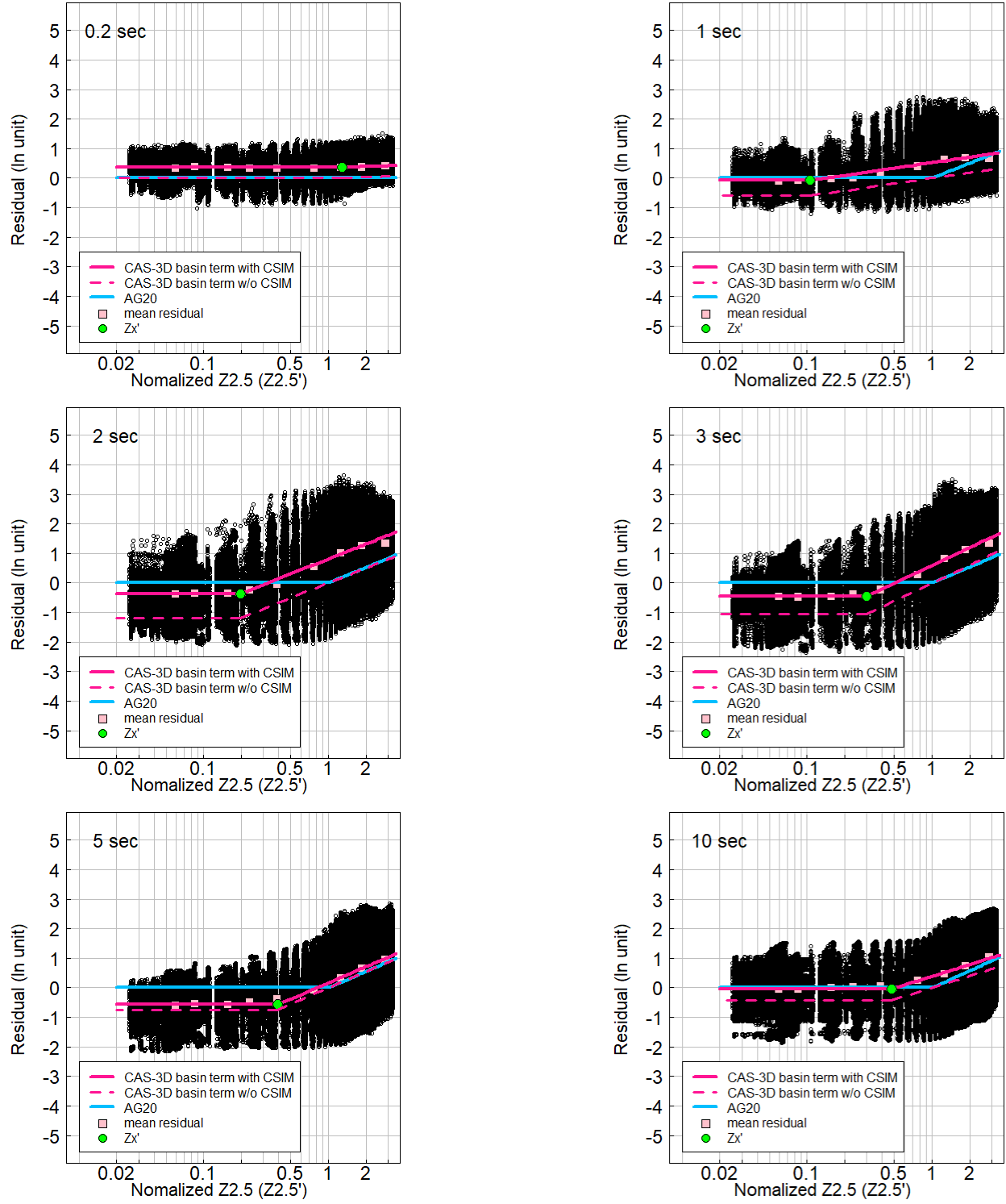


Figure 9: The updated basin-depth scaling ($Z'_{2.5}$ scaling) based on the 3-D simulation results.

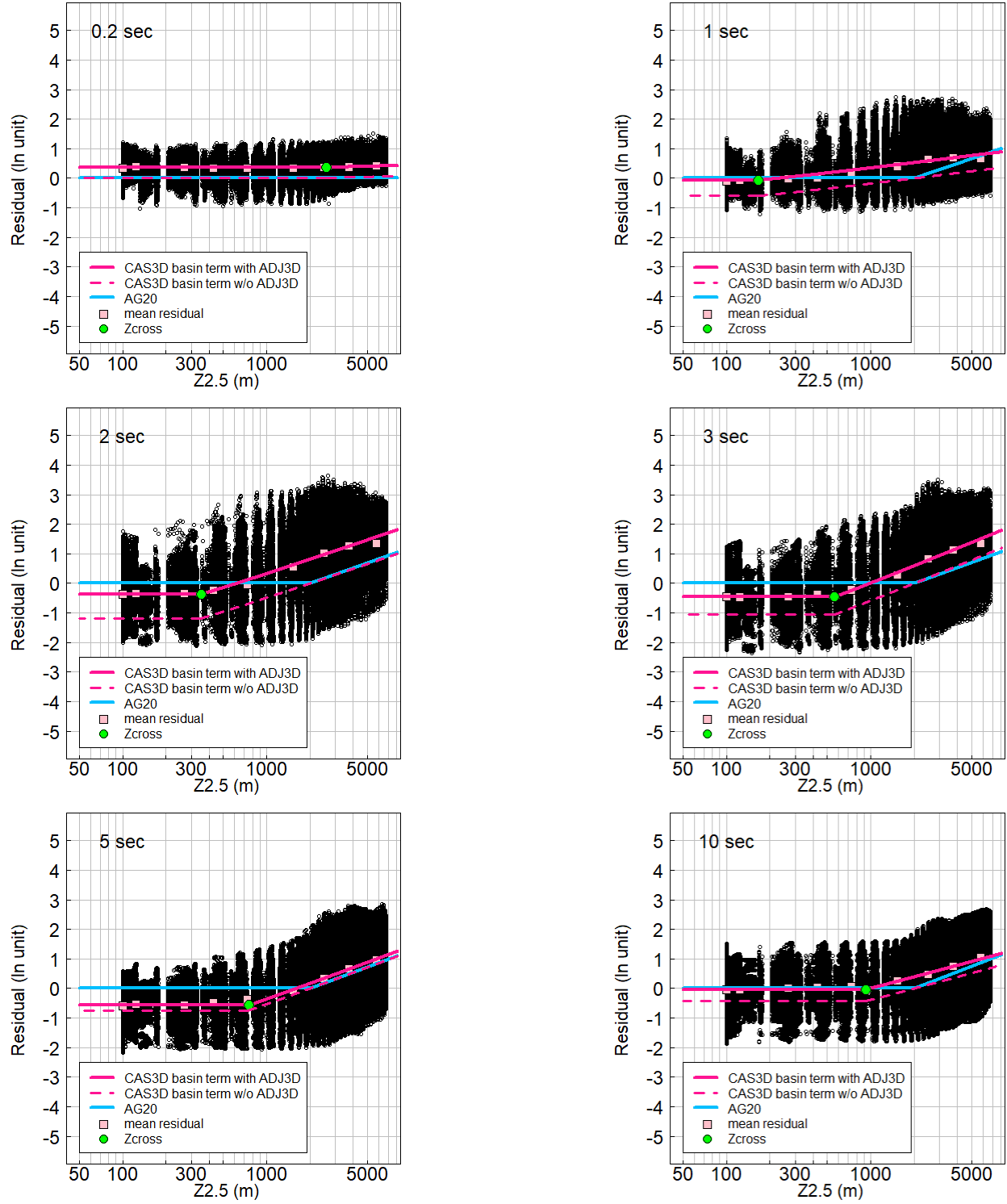


Figure 10: The updated basin-depth scaling based on the 3-D simulation results as a function of $Z_{2.5}$ without the normalization.

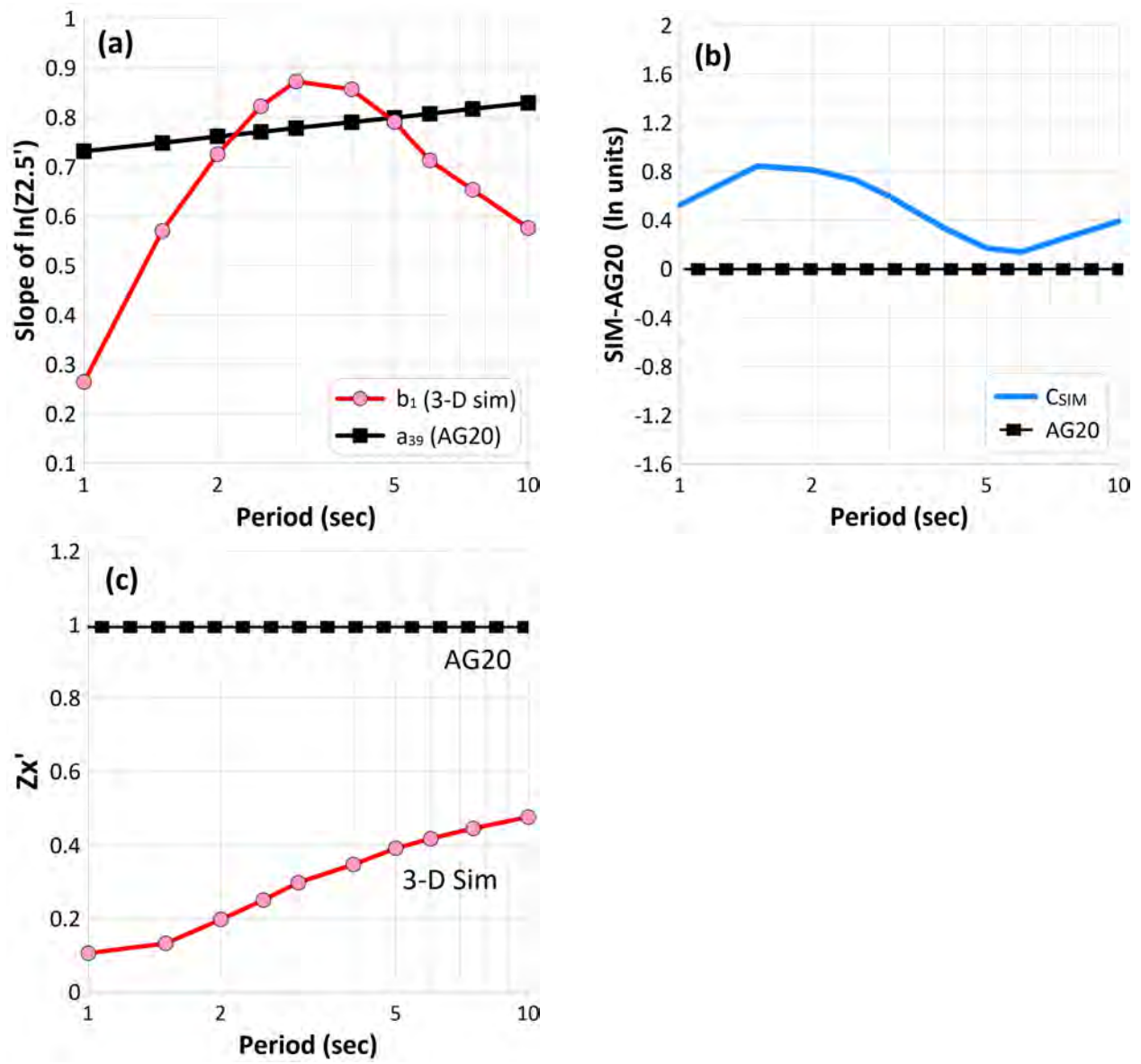


Figure 11: Coefficients for the AG20 and new basin-depth terms per period.

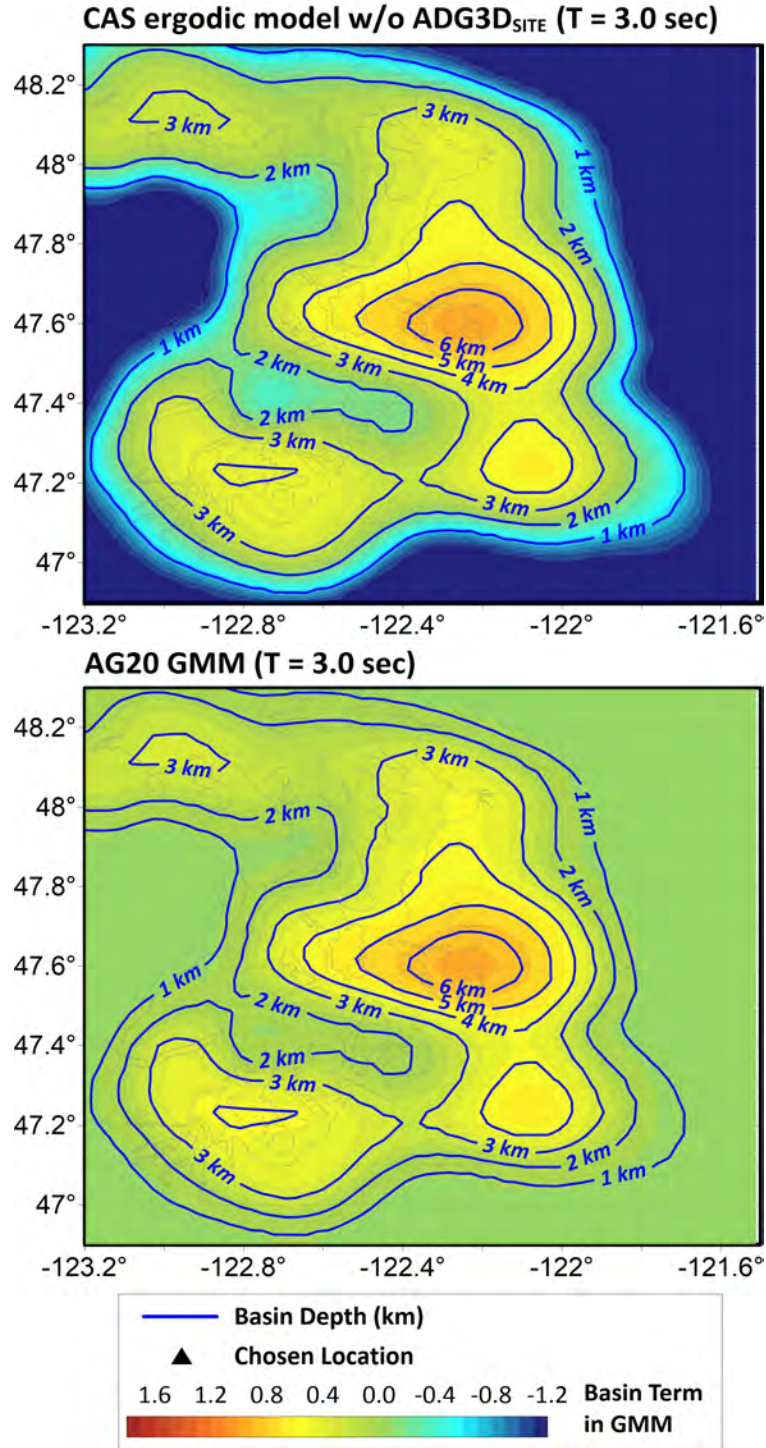


Figure 12: Maps of the $T=3$ sec basin depth term. The bottom frame is for the basin scaling in the original AG20 GMM. The top frame is for the basin scaling in the modified AG20 GMM based on the basin scaling in the 3-D simulations. The lines denote the 1-km contour interval for $Z_{2.5}$.

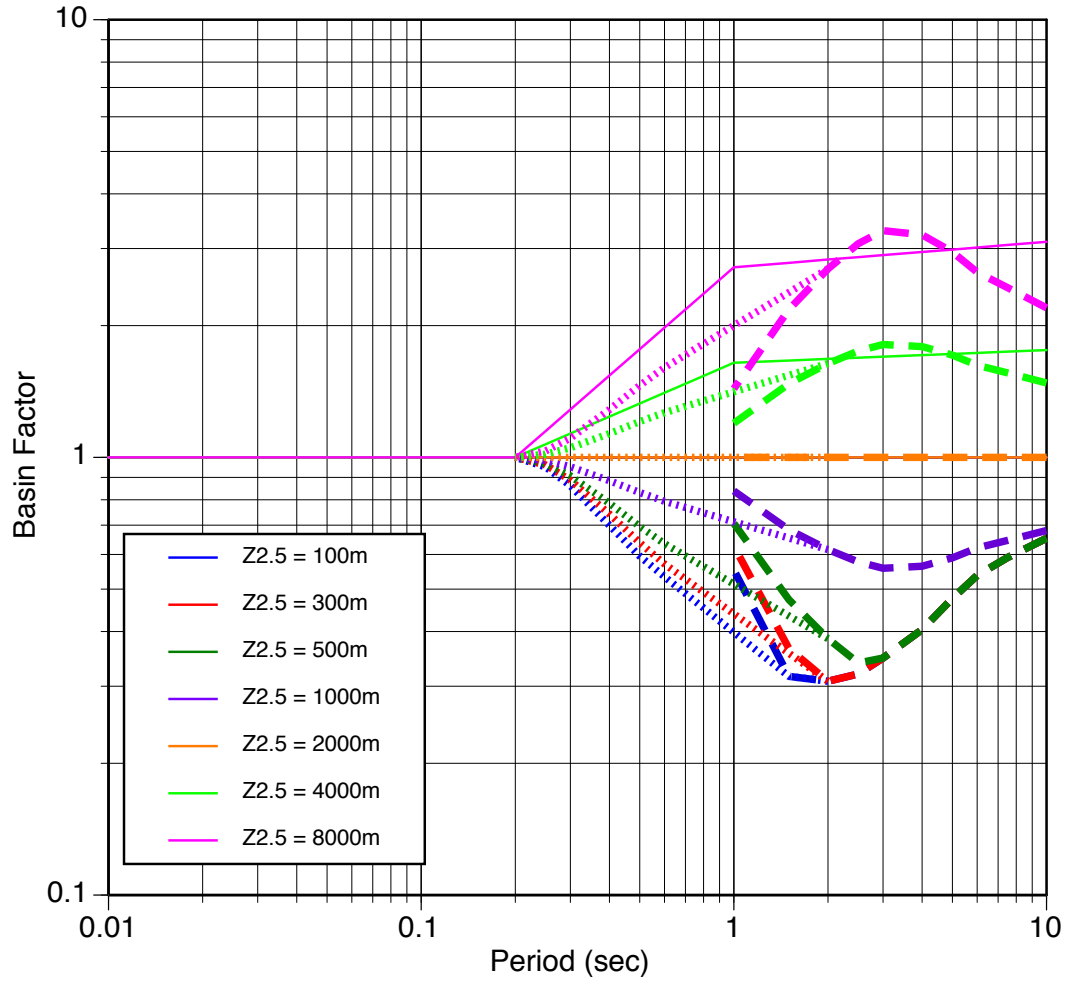


Figure 13: Comparison of the basin factor for $Z_{2.5}$ between 100 m and 8000 m. The thin solid lines show the scaling from the unmodified AG20 model. The long-dashed lines show the scaling from the M9 simulations. The extrapolation of the basin scaling from the M9 simulations to periods less than $T=2$ sec is shown by short dashed lines.

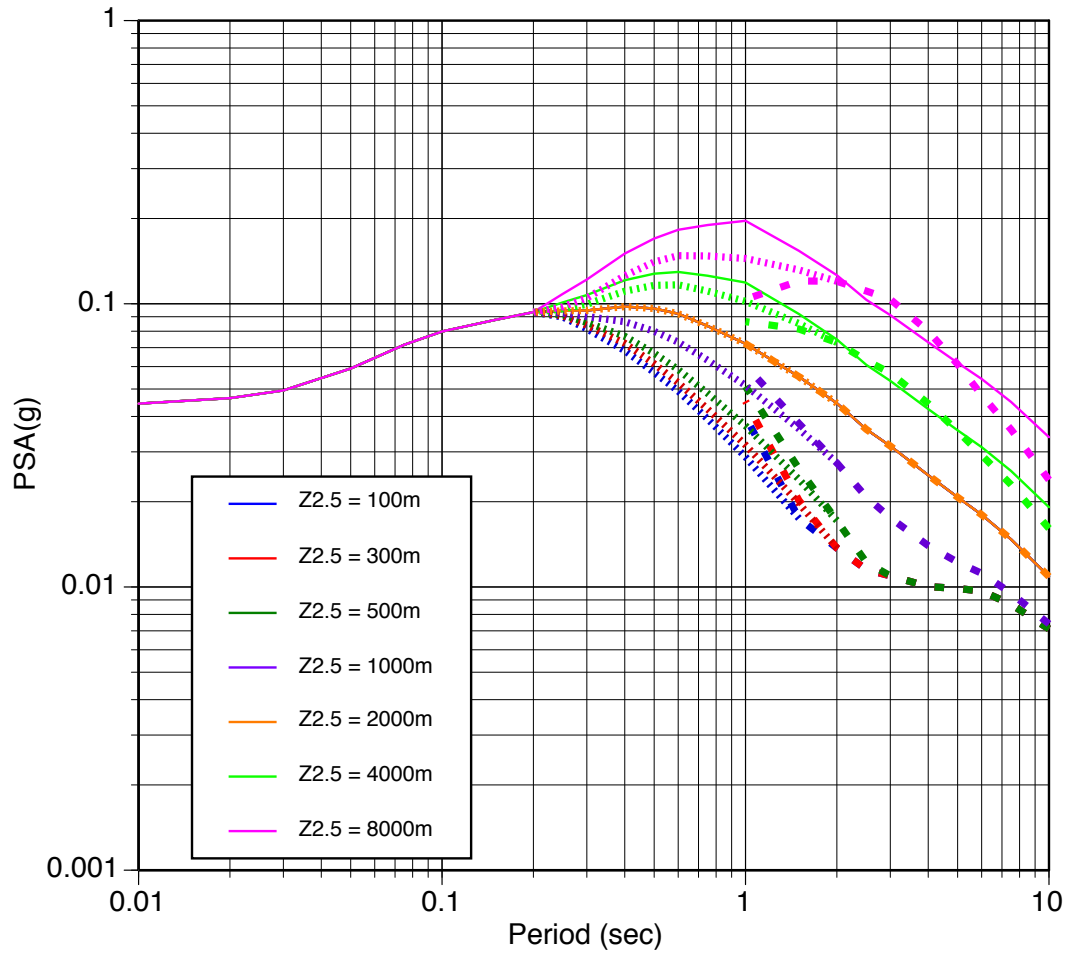


Figure 14: Comparison of the response spectra for M9 interface earthquake at $R=100$ km for the AG20 GMM using the unmodified basin scaling (thin solid lines), the M9 simulation basin scaling (long-dashed lines), and using the M9 scaling extrapolated to short periods (short-dashed lines).

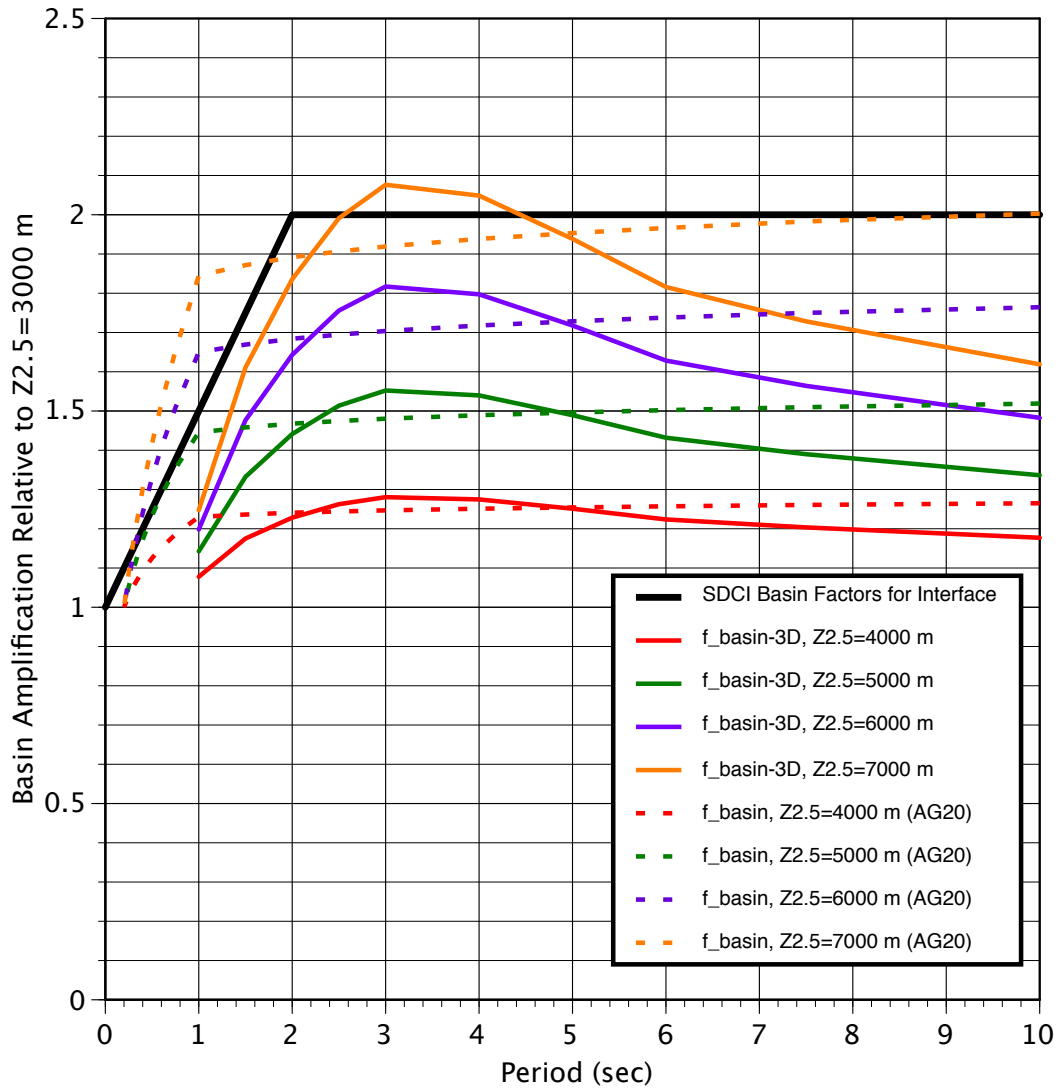


Figure 15: Comparison of the normalized basin factors from AG20 and from the 3-D simulations with the basin factors used by the SDCI.

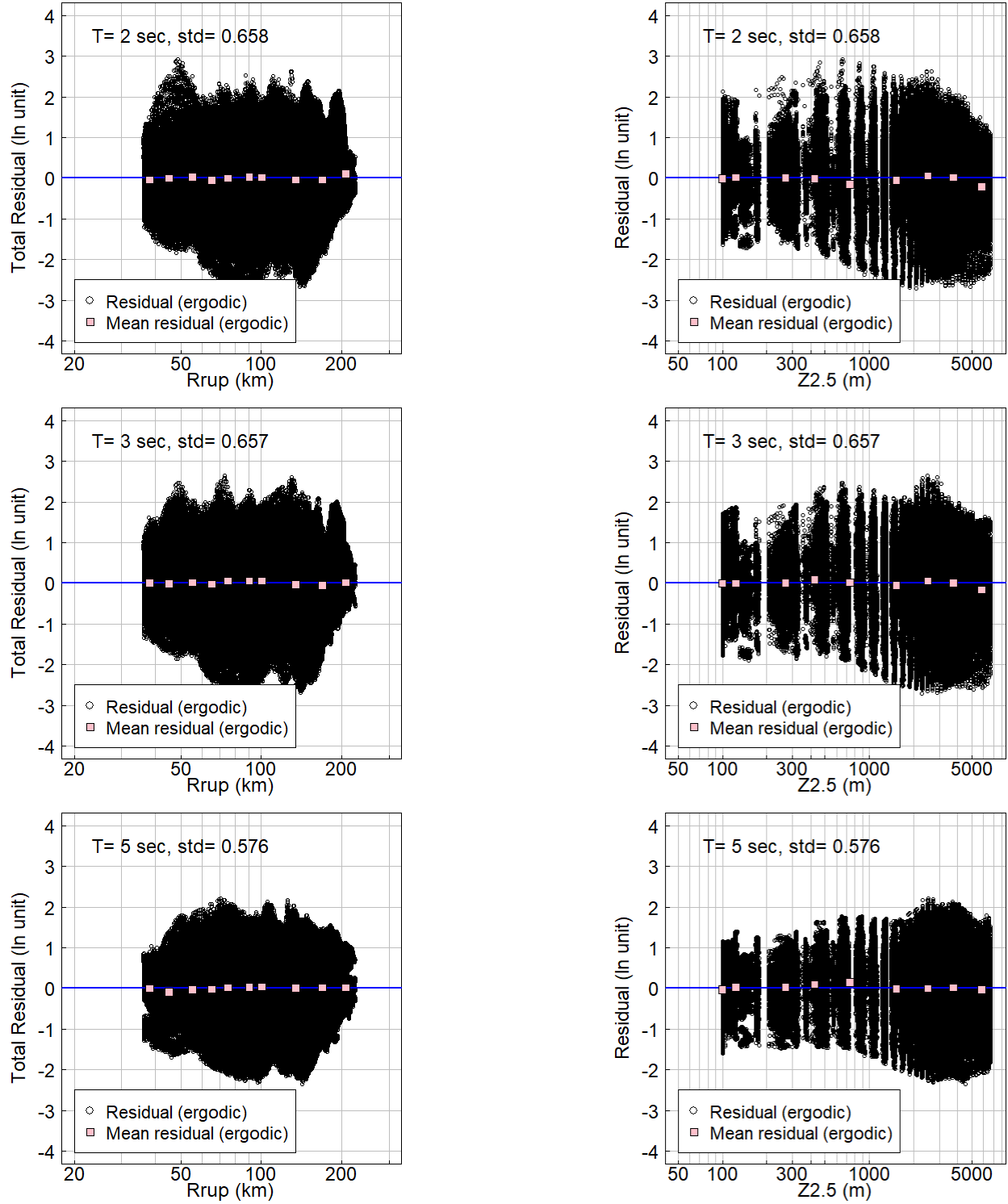


Figure 16: Total Residuals of the ergodic model versus Rrup (*left*) and Z2.5 (*right*) for $T = 1.5, 3$ and 5 sec.

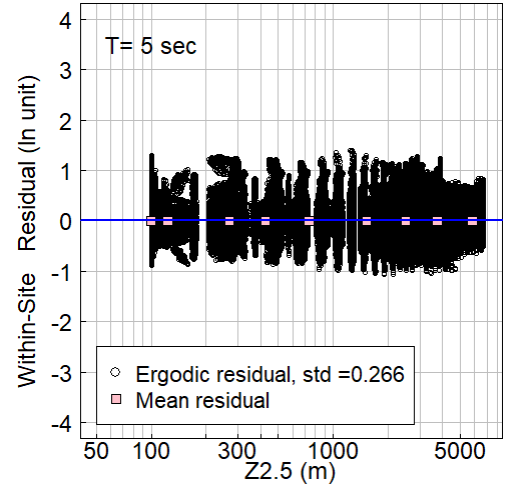
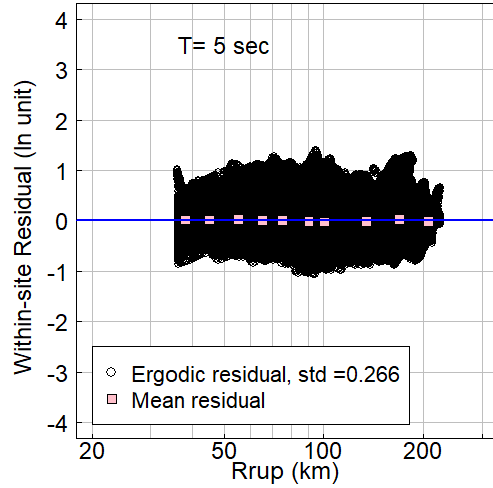
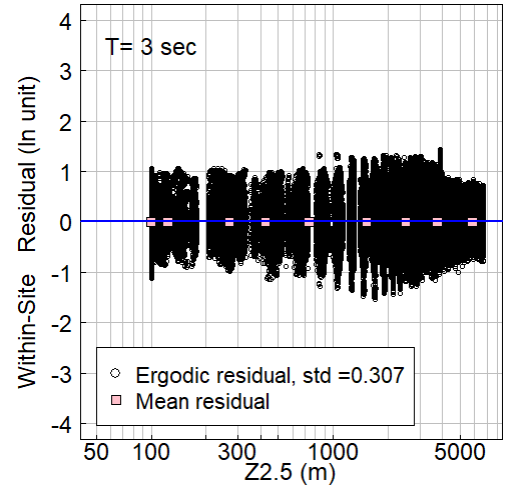
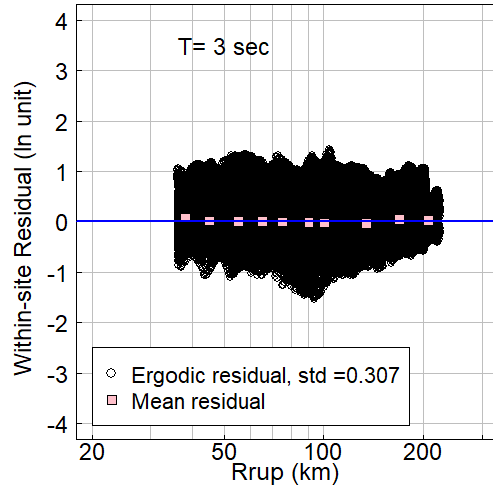
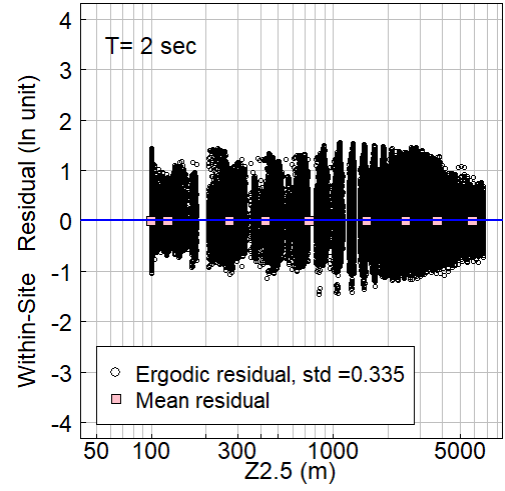
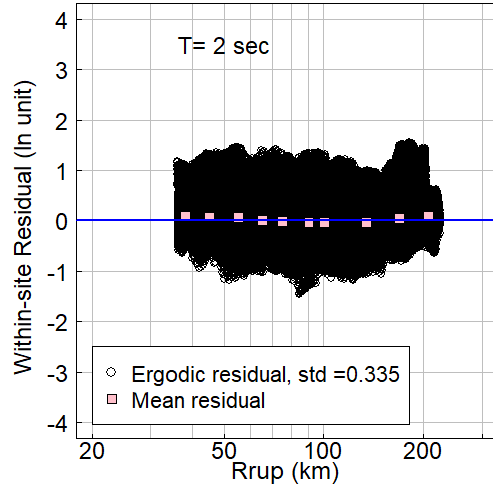


Figure 17: Within-site Residuals of the ergodic model versus Rrup (*left*) and Z2.5 (*right*) for $T = 1.5, 3$ and 5 sec.

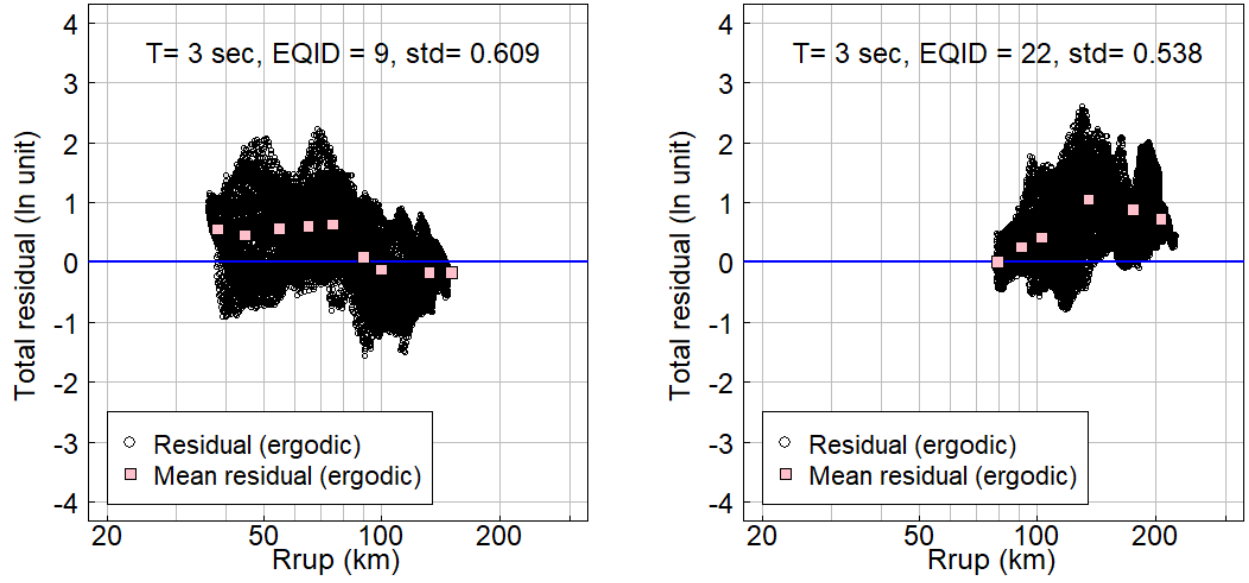


Figure 18: 3sec-total-residuals of the adjusted AG20 GMM (ergodic model) versus R_{RUP} scaling for two target simulations (ID9 and ID22 in Figure 3).

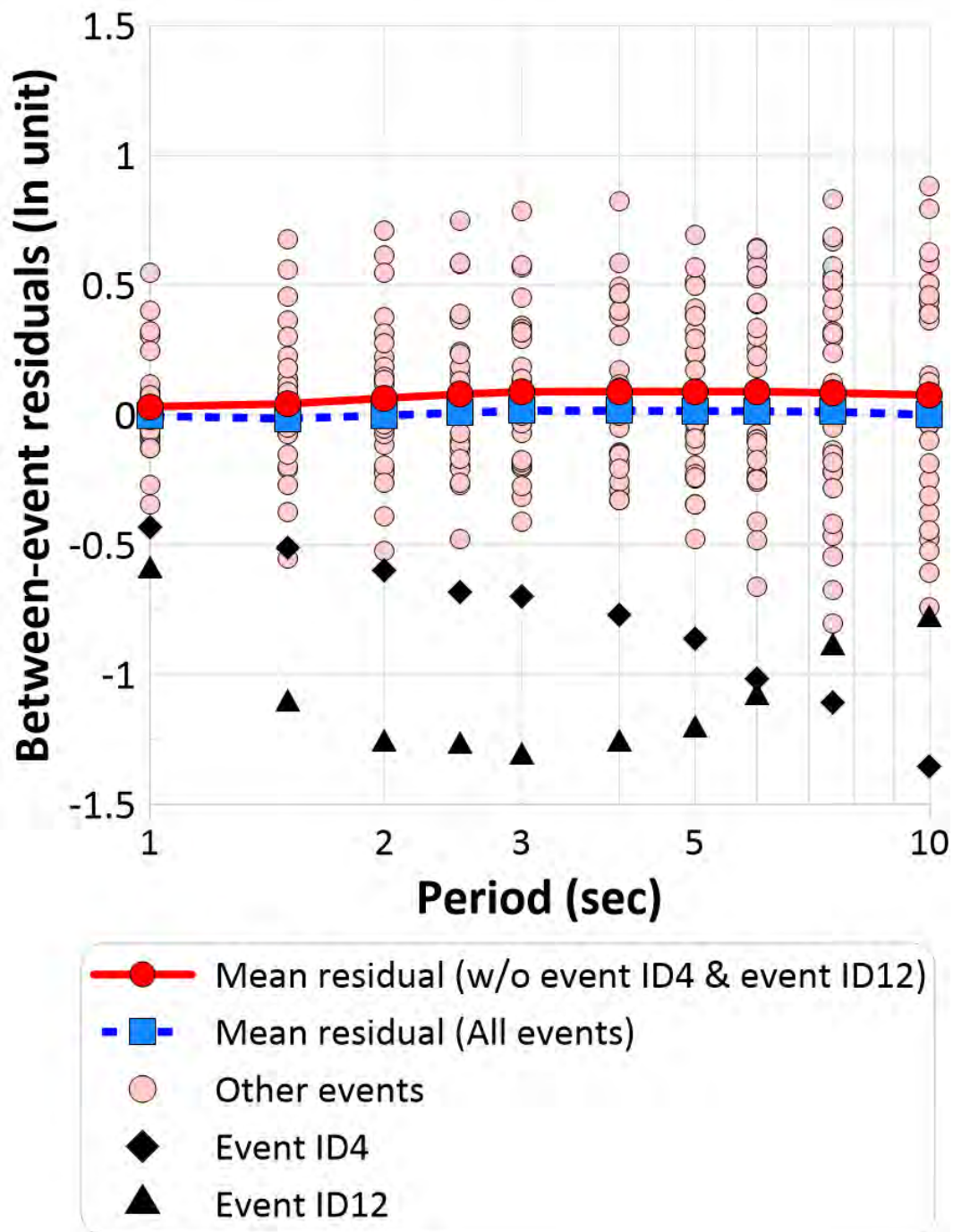


Figure 19: Between-event residuals of the ergodic model versus period.

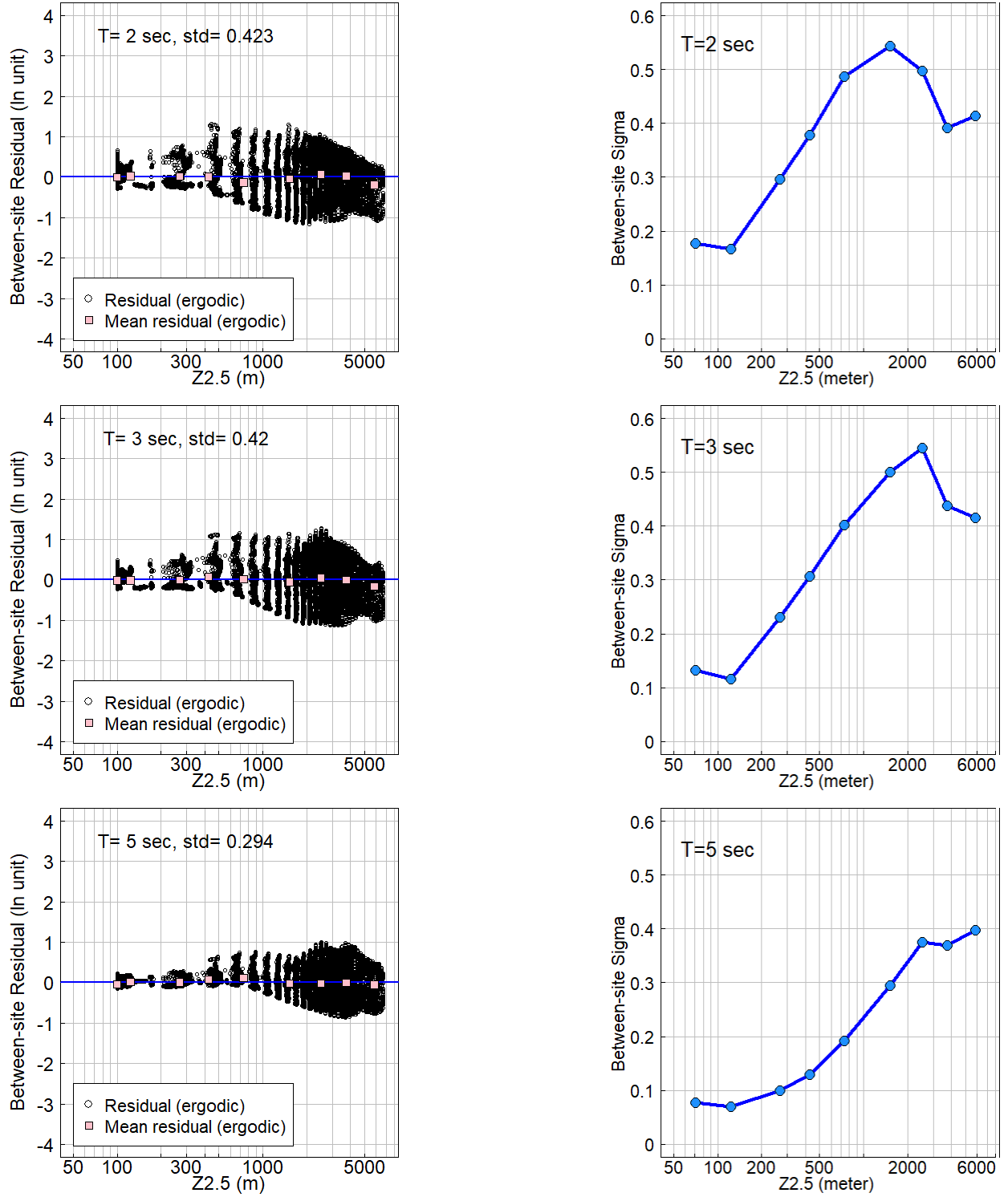


Figure 20: *left*: Between-site residuals of the ergodic model versus $Z_{2.5}$ scaling for 2sec, 3 sec and 5 sec. *right*: Between-site sigma (ϕ_{S2S}) for various $Z_{2.5}$ ranges .

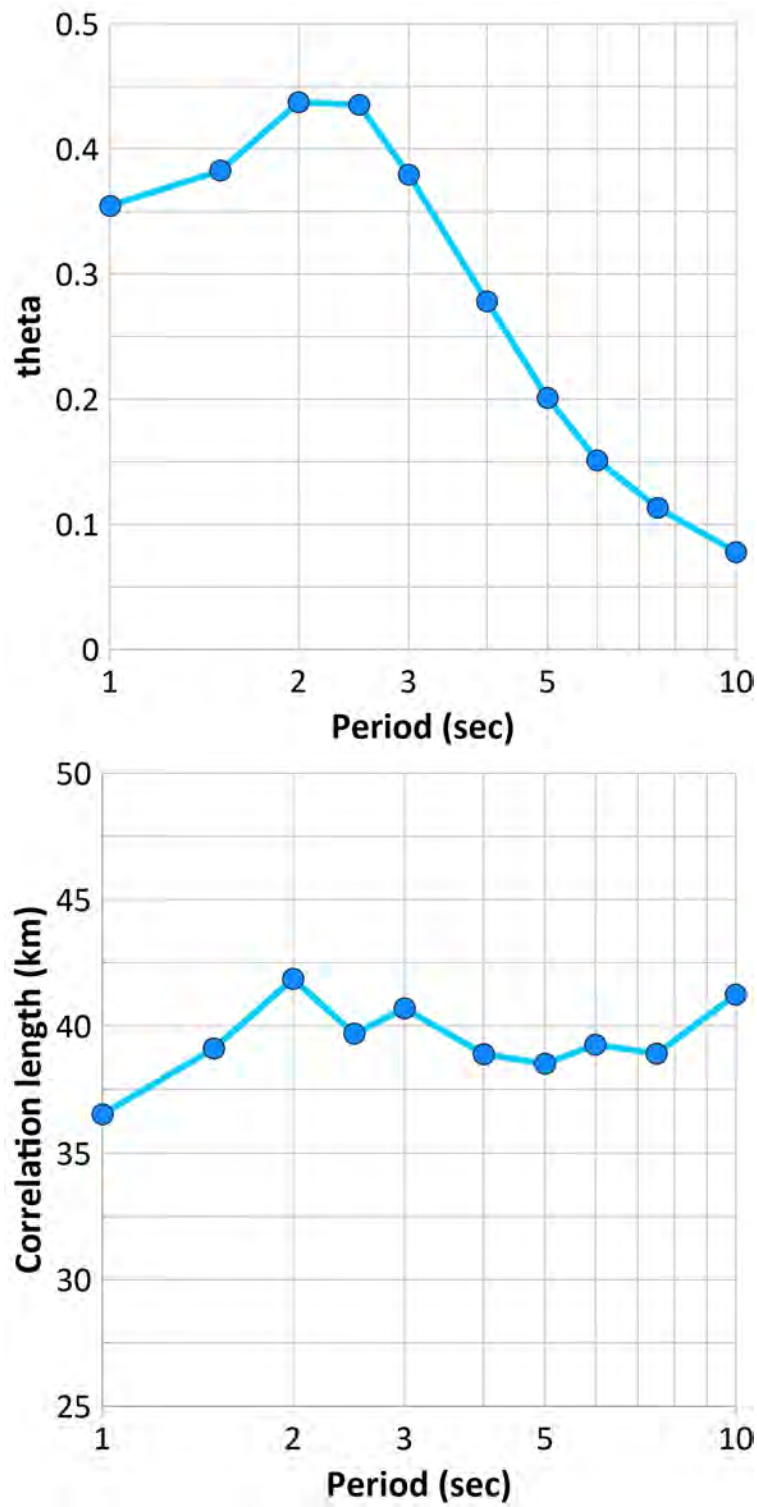


Figure 21: Hyperparameter of Matern covariance function for periods: variance, σ^2 (*left*), and correlation length, ρ (*right*).

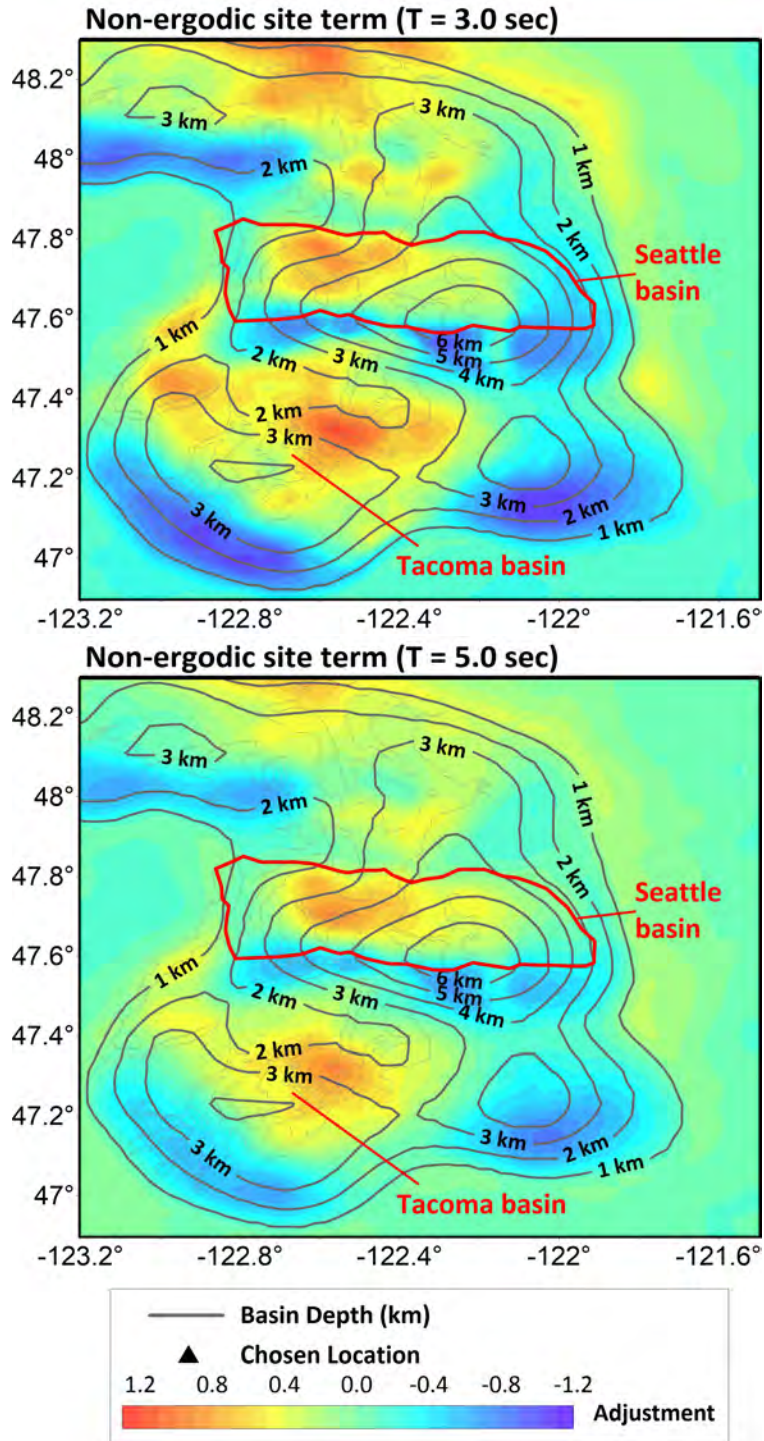


Figure 22: Spatially varying non-ergodic site term at $T = 3$ sec and 5 sec. This is based on the non-ergodic methodology with the spatial correlation of the site terms. Lines denote the $Z_{2.5}$ contour (1-km contour interval).

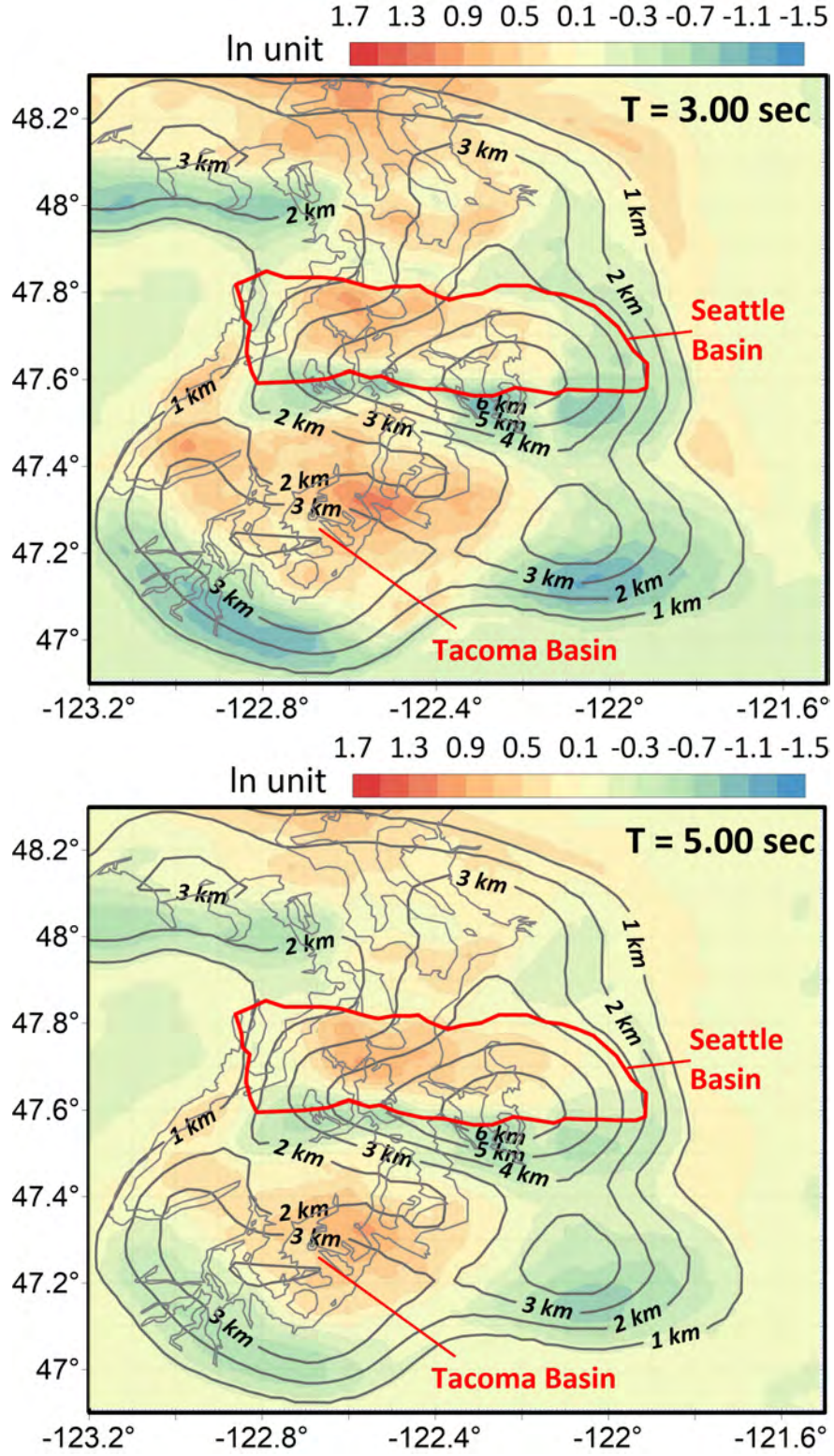


Figure 23: Maps of the between-site residual ($\delta S2S$) computed using random effects (without the spatial correlation) for $T = 3$ sec and 5 sec. Lines denote the $Z_{2.5}$ contour (1-km contour interval).

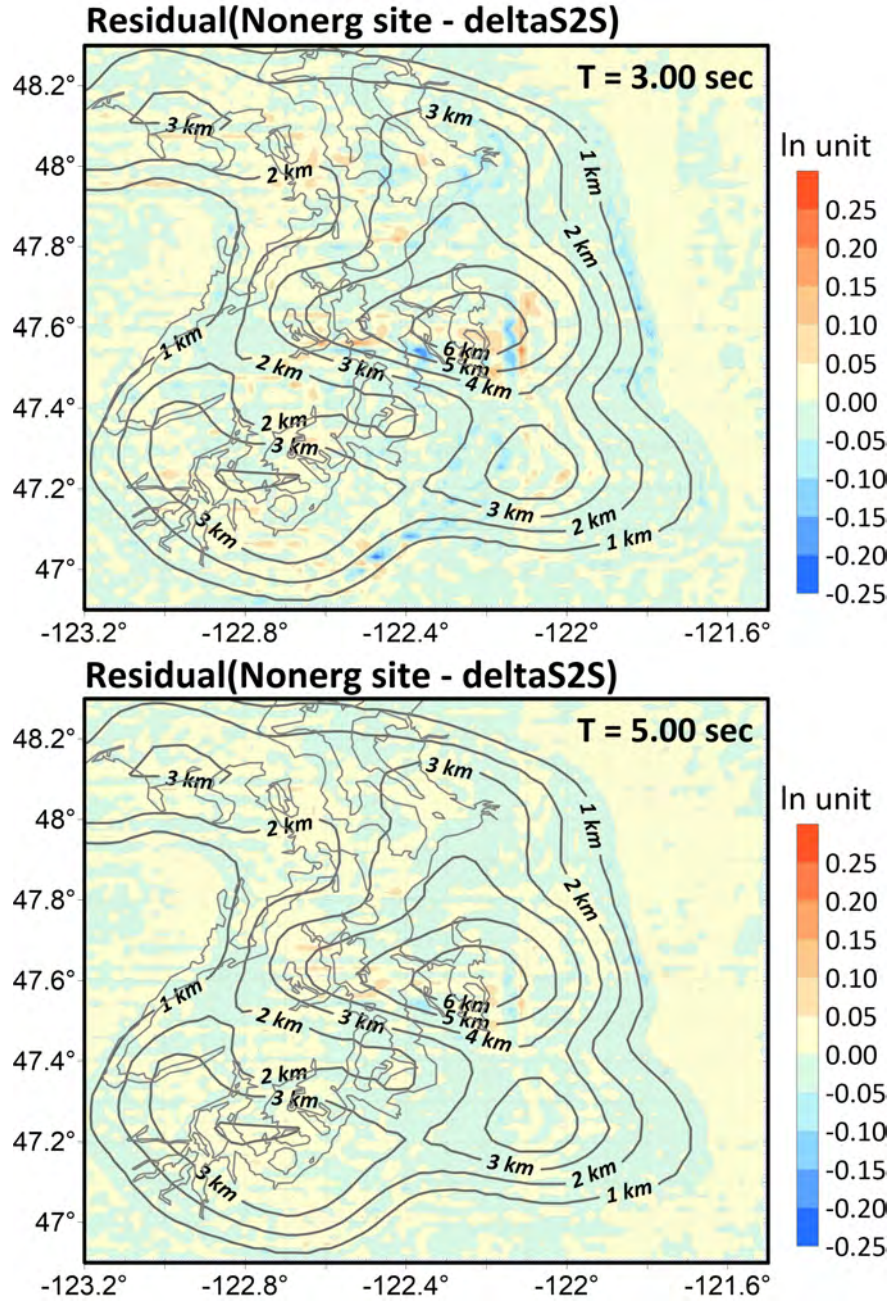


Figure 24: Maps of the difference in the between-site terms with and without spatial correlation ($\ln(Site_{nonerg}) - \ln(\delta S2S)$) for $T = 3$ sec and 5 sec. Lines denote the $Z_{2.5}$ contour (1-km contour interval).

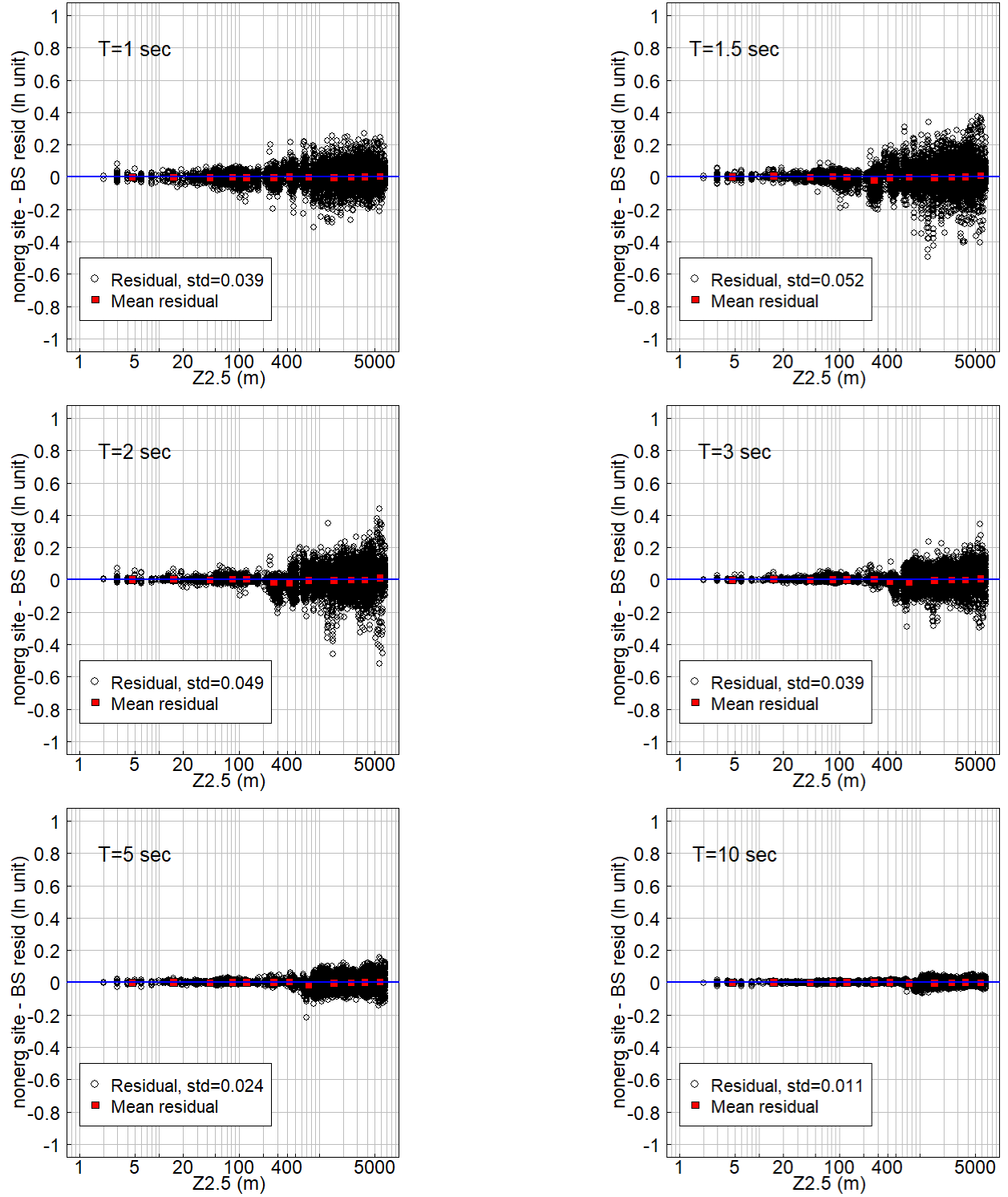


Figure 25: Difference between non-ergodic site terms and the ergodic between-site residuals ($\delta S2S_{nonerg} - \delta S2S$) and site terms for 1sec to 10 sec.

Residual (nonerg site –
delta S2S), T=2.0

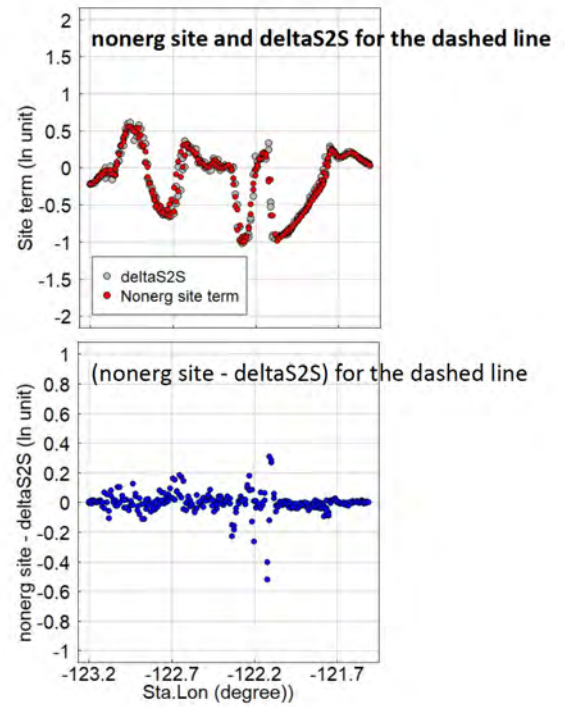
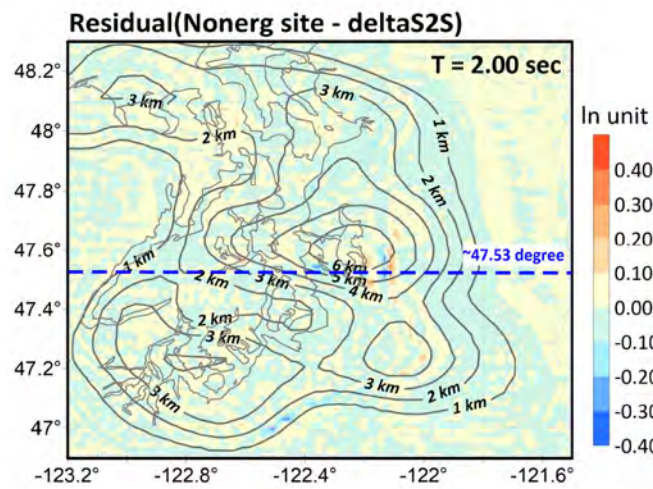


Figure 26: Cross section of the difference between non-ergodic site terms and the ergodic between-site residuals ($\delta S2S_{nonerg} - \delta_{S2S}$) and site terms for T=2sec.

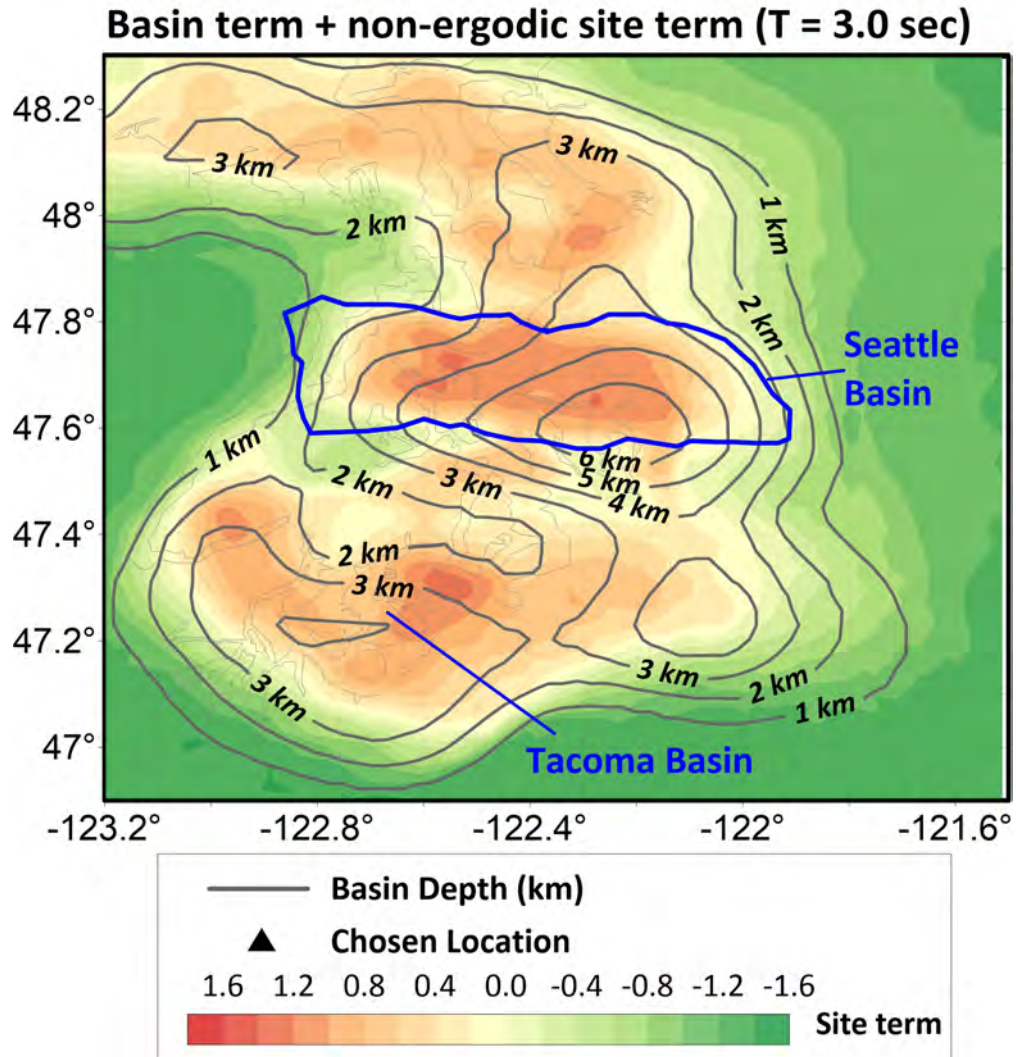


Figure 27: Total site term (Basin-depth term + non-ergodic site term) for $T = 3$ sec. Lines denote the $Z_{2.5}$ contour (1-km contour interval).

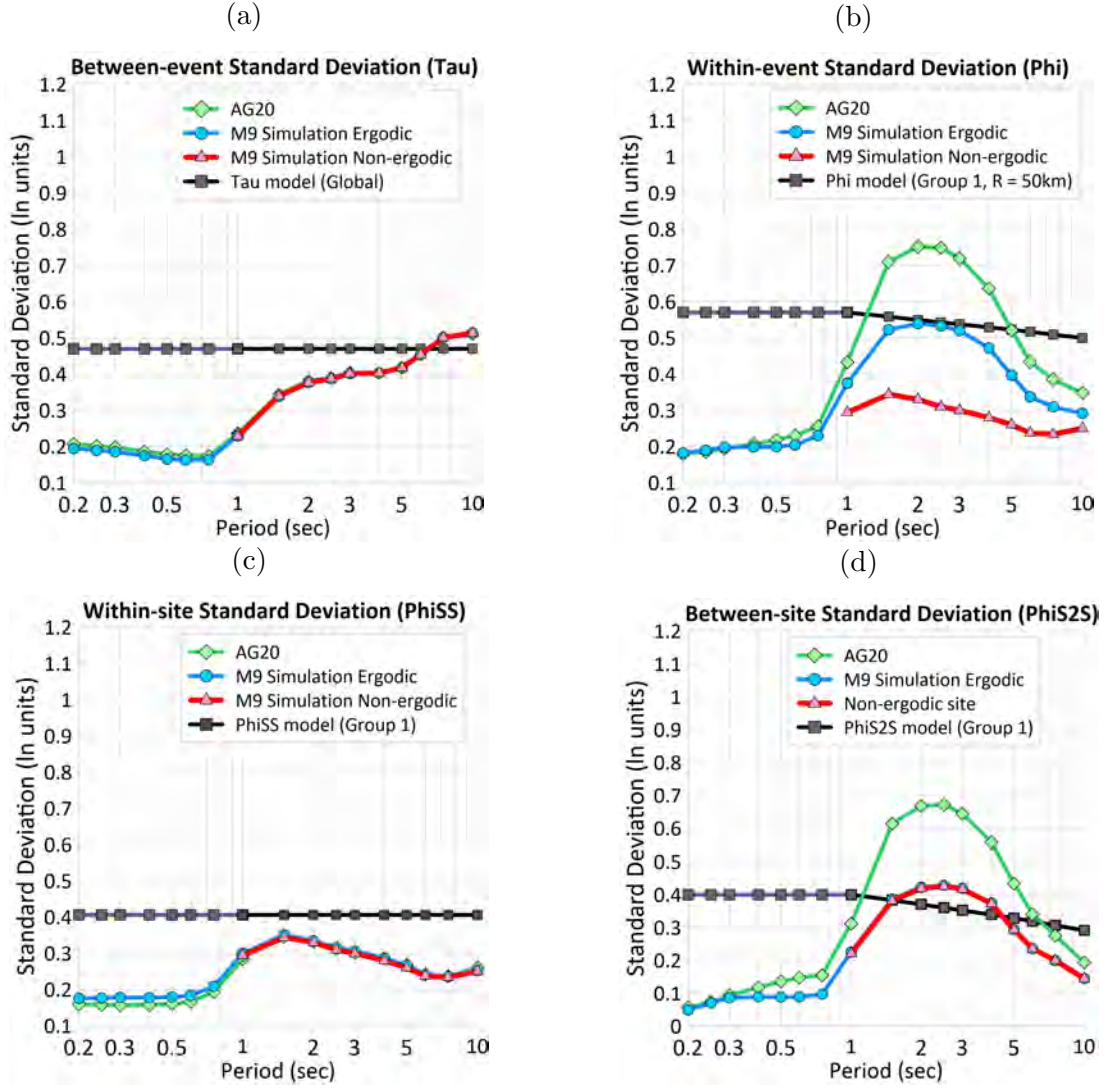


Figure 28: (a) Between-event standard deviation (τ). (b) Within-event standard deviation (ϕ). (c) Single-station within-event standard deviation (ϕ_{SS}). (d) Between-site standard deviation (ϕ_{S2S}).

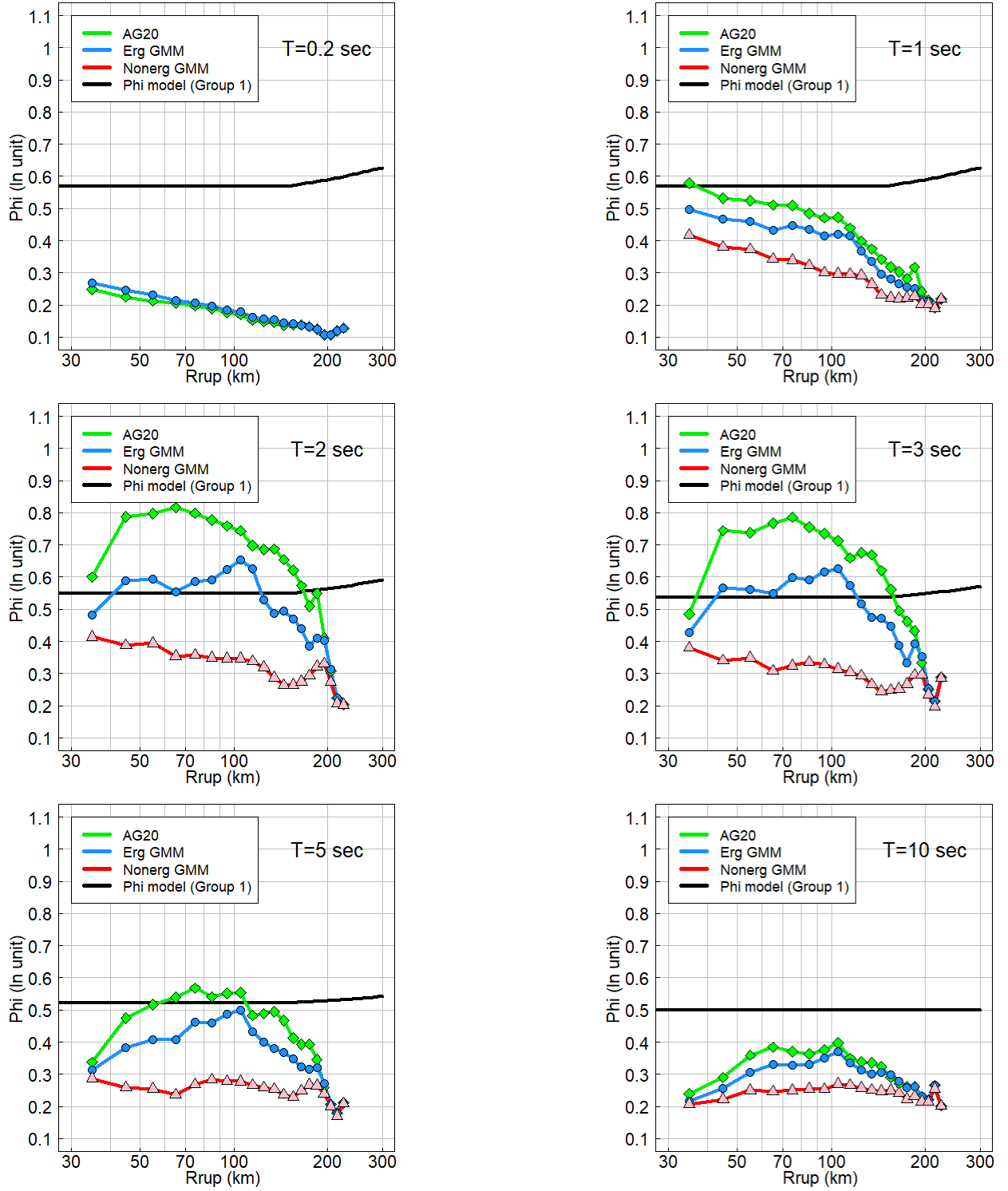


Figure 29: ϕ and ϕ_{SS} versus R_{rup} for $T = 0.2, 1, 2, 3, 5$ and 10 sec. The red curves are for ϕ_{SS} and the other curves are for ϕ . The black curves show the Group 1 ϕ model from AG20 used for Cascadia.

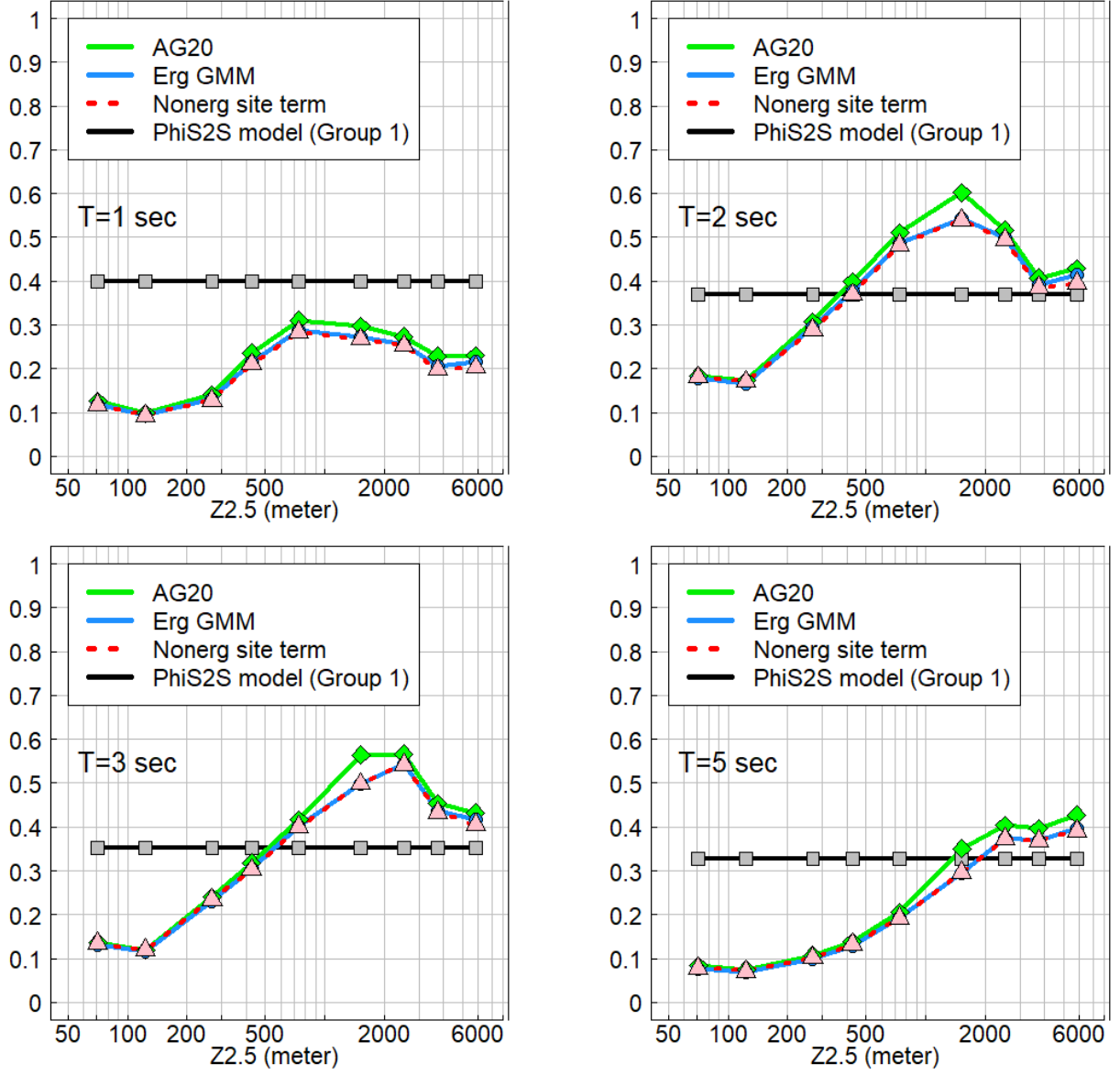


Figure 30: ϕ_{S2S} versus $Z_{2.5}$ for $T = 1, 2, 3$ and 5 sec. The black curves show the Group 1 ϕ_{S2S} model from AG20 used for Cascadia.

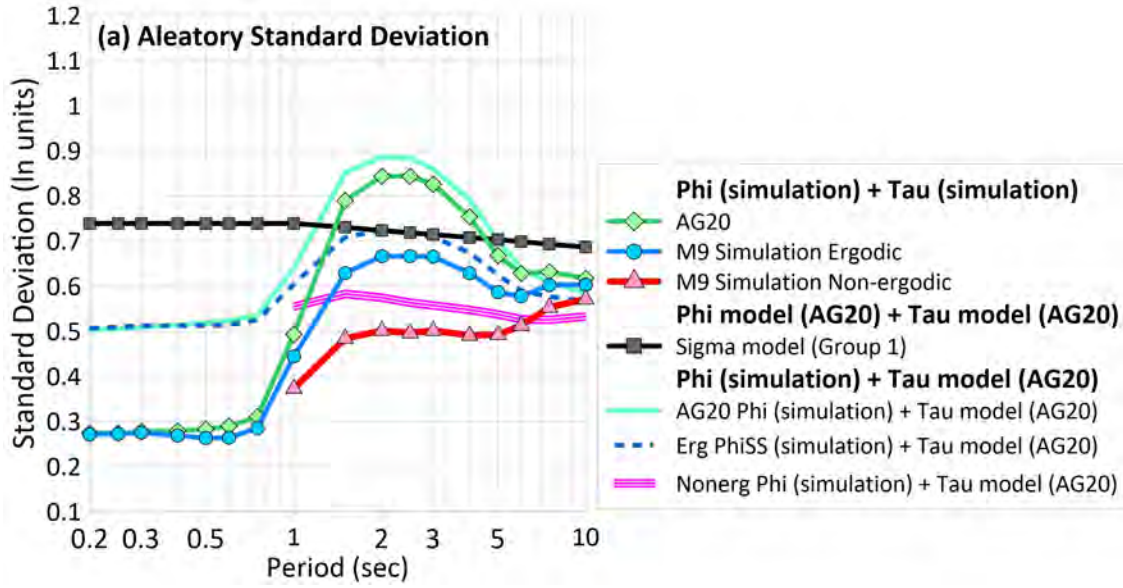


Figure 31: The total standard deviation values are shown for three combinations of the ϕ and τ : the first set uses the ϕ and τ from the simulations; the second set uses the empirical ϕ and τ from the AG20 GMM; and the third set uses the ϕ from the simulations and the global τ from the AG20 GMM. The first set includes the AG20 GMM (green curve), modified AG20 ergodic GMM (blue curve), and non-ergodic GMM (red curve). The second set is the sigma from the AG20 GMM (black curve). The third set includes AG20 GMM (light blue curve), modified AG20 ergodic GMM (dashed blue curve), and non-ergodic GMM (pink curve).

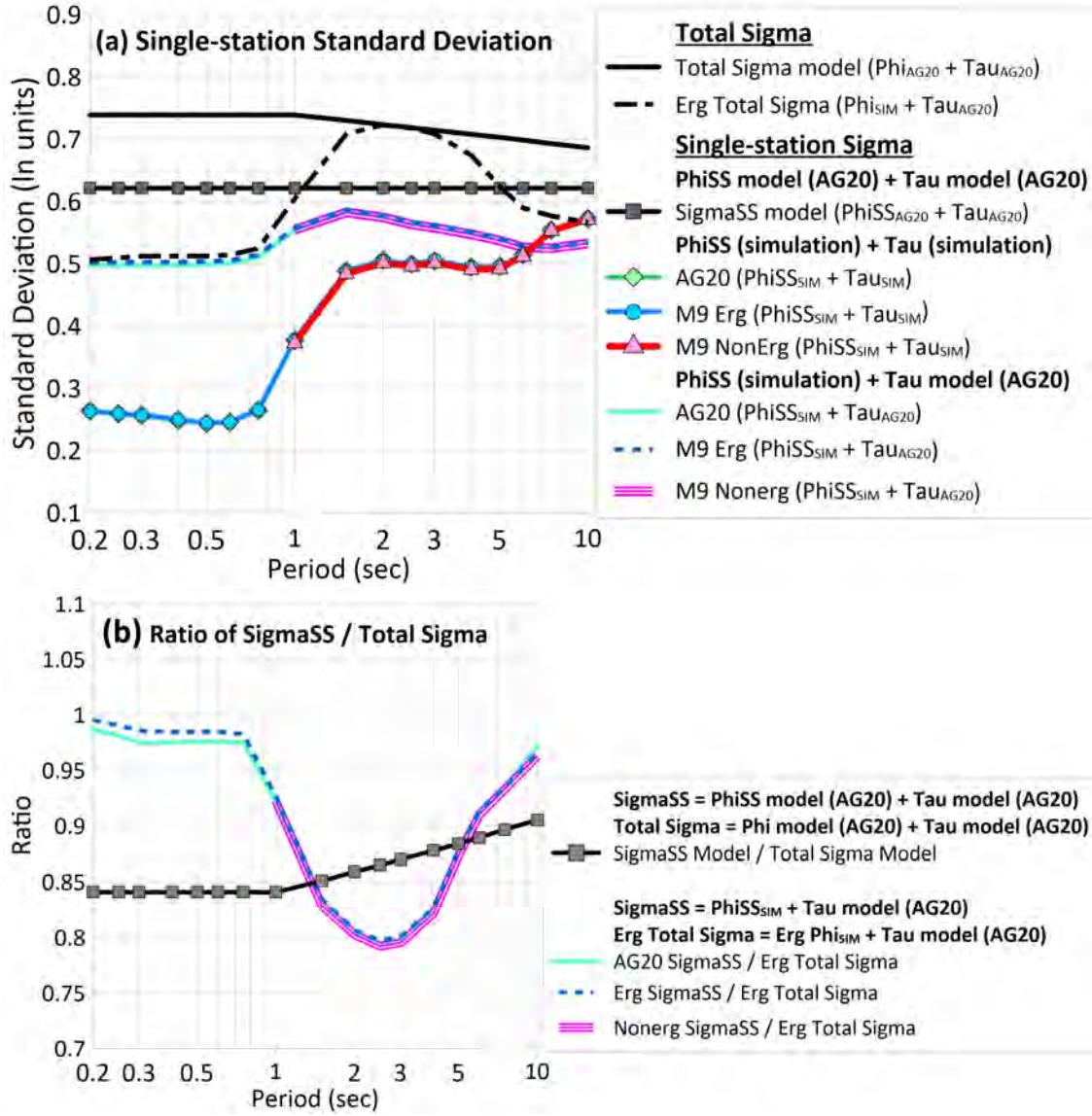


Figure 32: (Top) Comparison of single-station standard deviation (σ_{SS}) models. The σ_{SS} model from AG20 is shown by the black squares. For the simulation-based approaches, the σ_{SS} models are computed using the τ from the simulations (lower curves) and using the τ from AG20 (upper curves). (Bottom) Ratio of σ_{SS} to the total σ from AG20 for Cascadia.

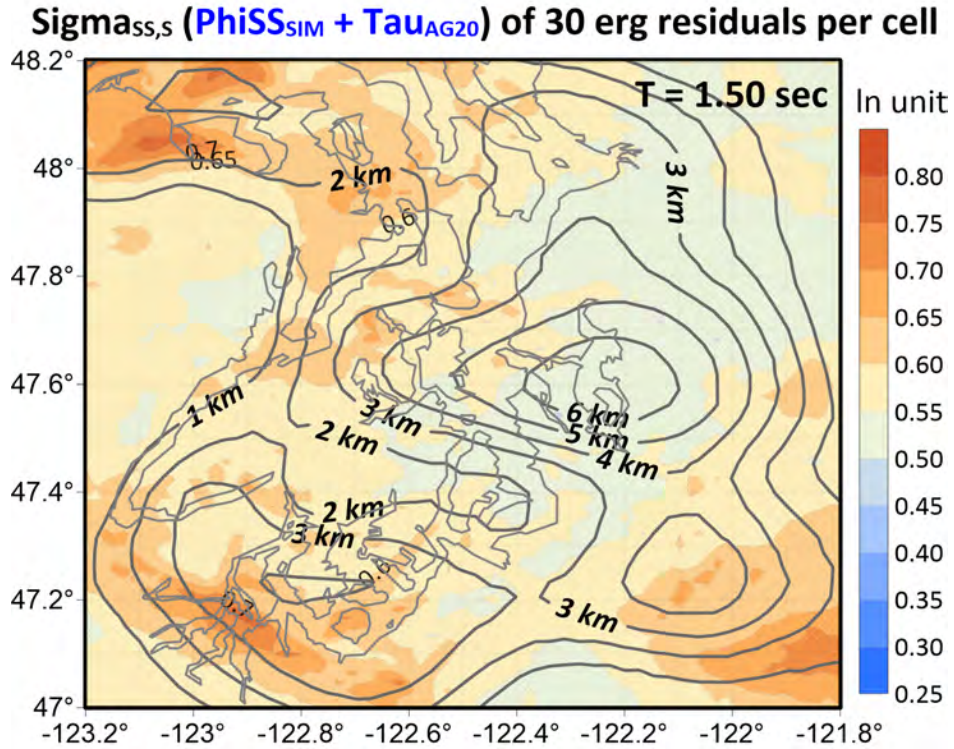
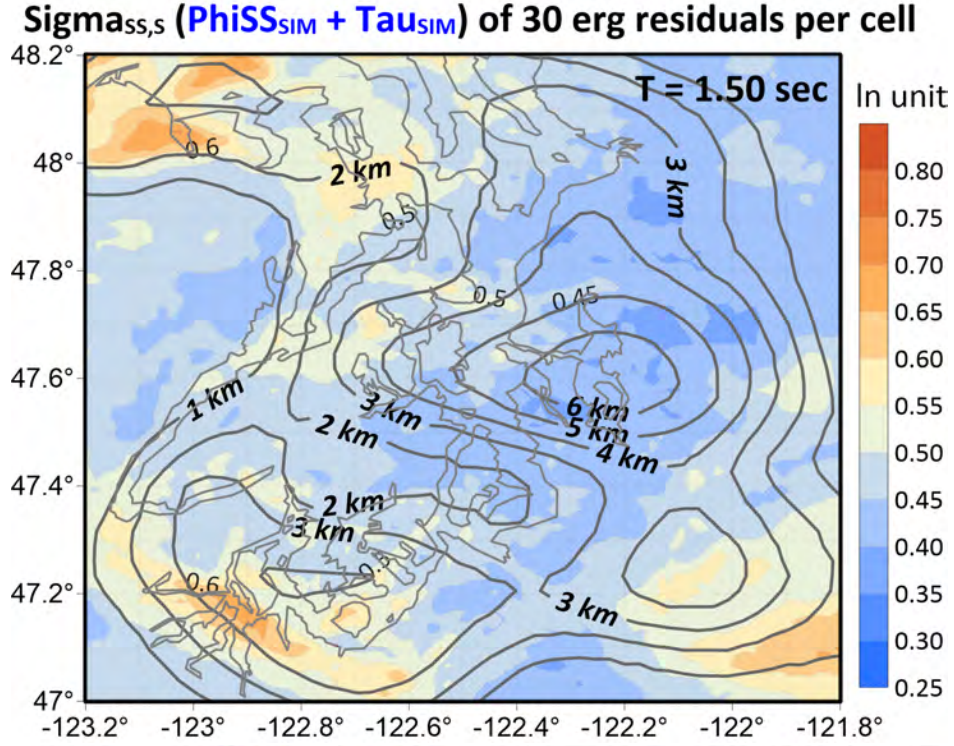


Figure 33: Maps of single-station sigma ($\sigma_{SS,S}$) for $T = 1.5 \text{ sec}$. The top frame uses the τ from the simulations and the bottom frame used τ from AG20.

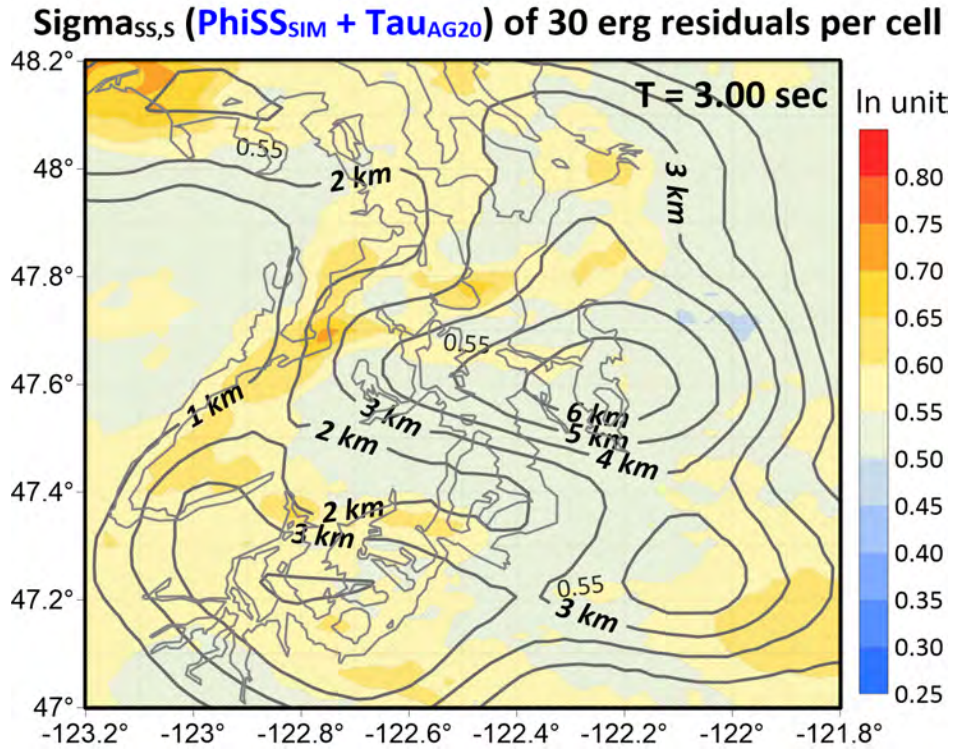
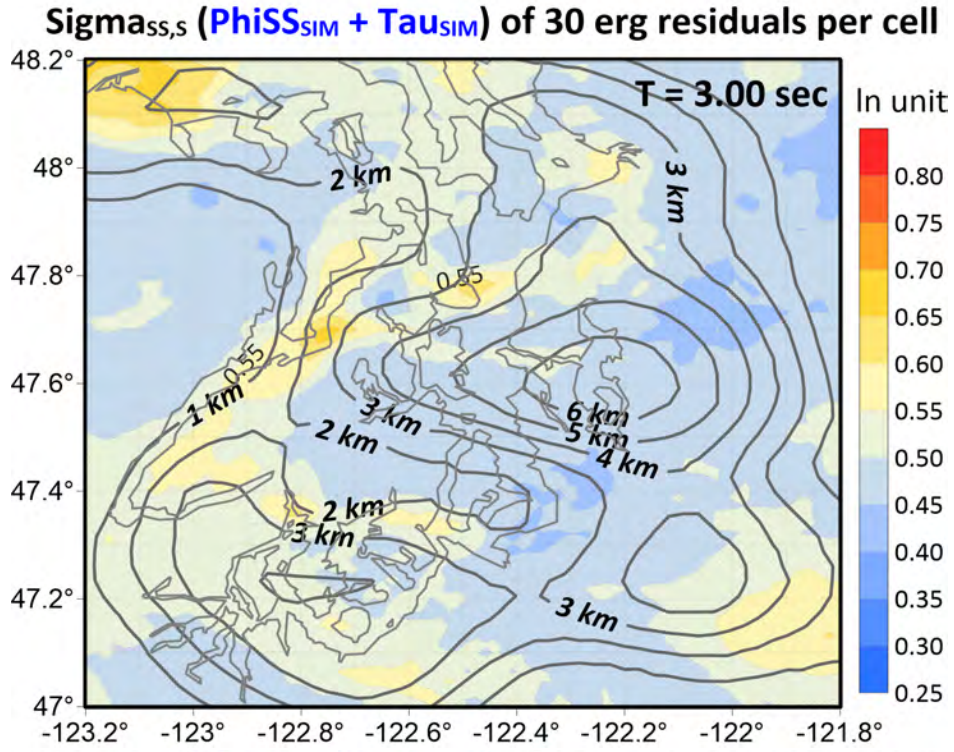


Figure 34: Maps of single-station sigma ($\sigma_{ss,s}$) for $T = 3$ sec. The top frame uses the τ from the simulations and the bottom frame used τ from AG20.

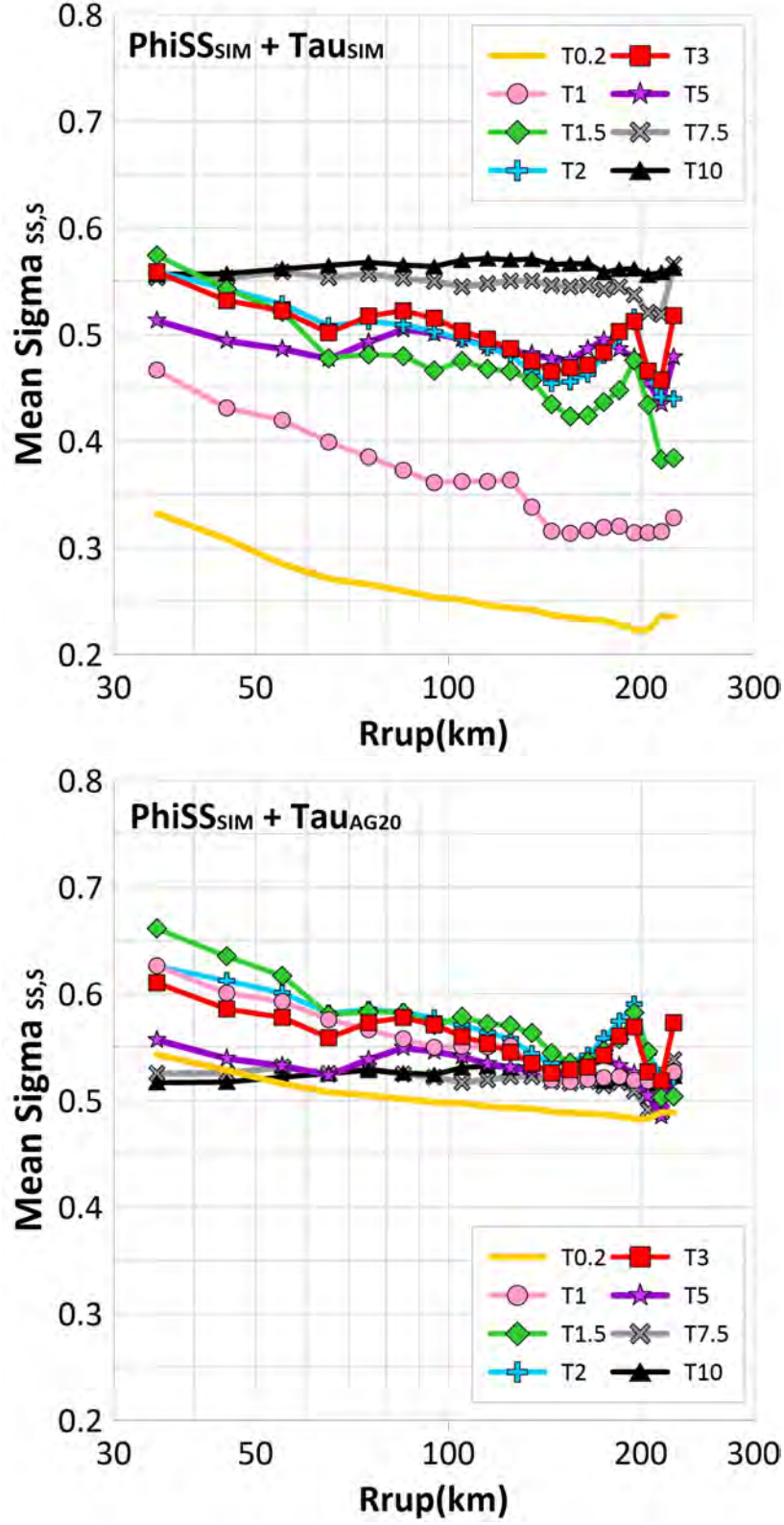


Figure 35: Mean $\sigma_{SS,S}$ for distance bins using the τ from the simulations (upper) and the τ from AG20 (bottom).

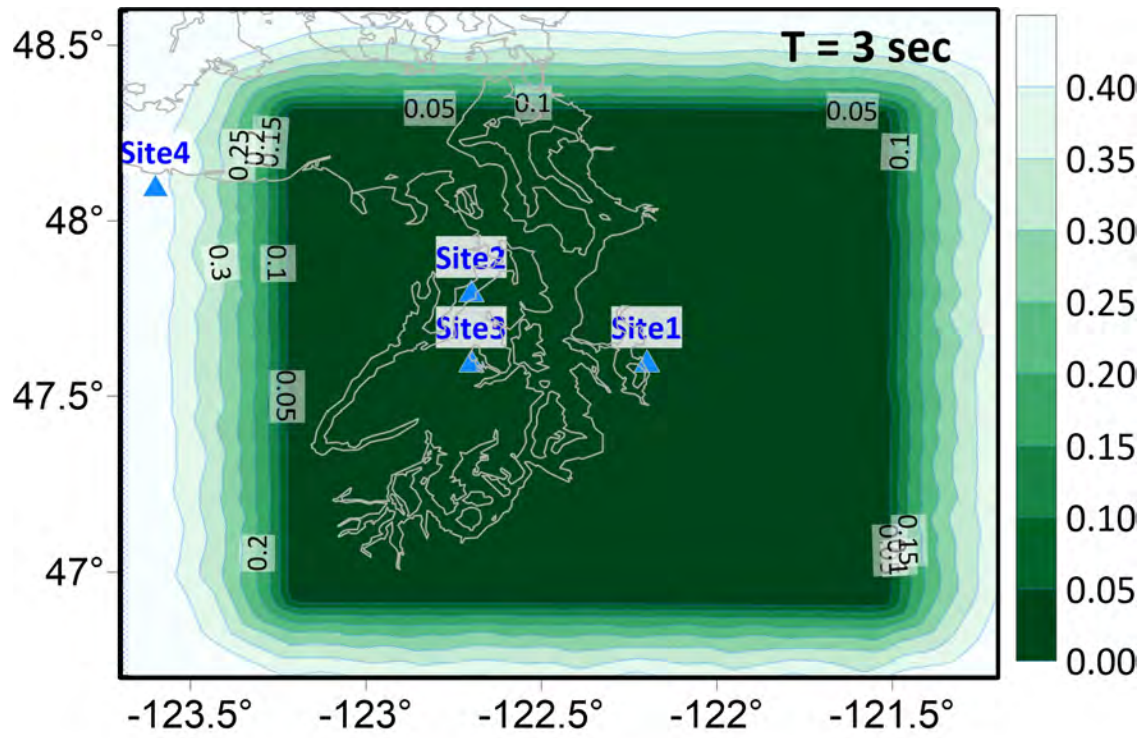


Figure 36: Map of the epistemic uncertainty of the non-ergodic site term ($\sigma_{\mu-SIM}$) for $T = 3$ sec.

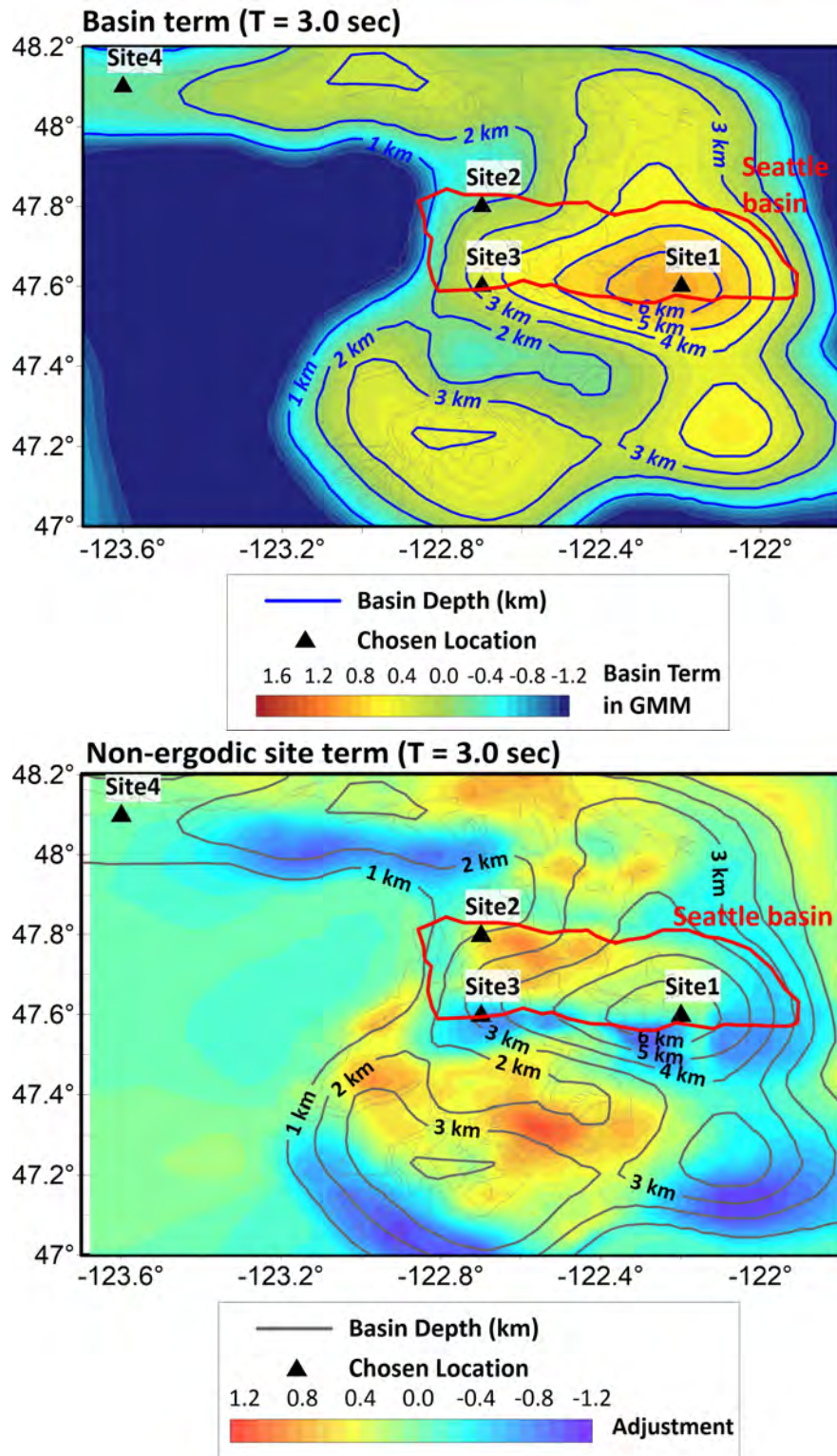


Figure 37: Location of the four example site in terms of the basin-depth and the non-ergodic site term at T=3 sec.

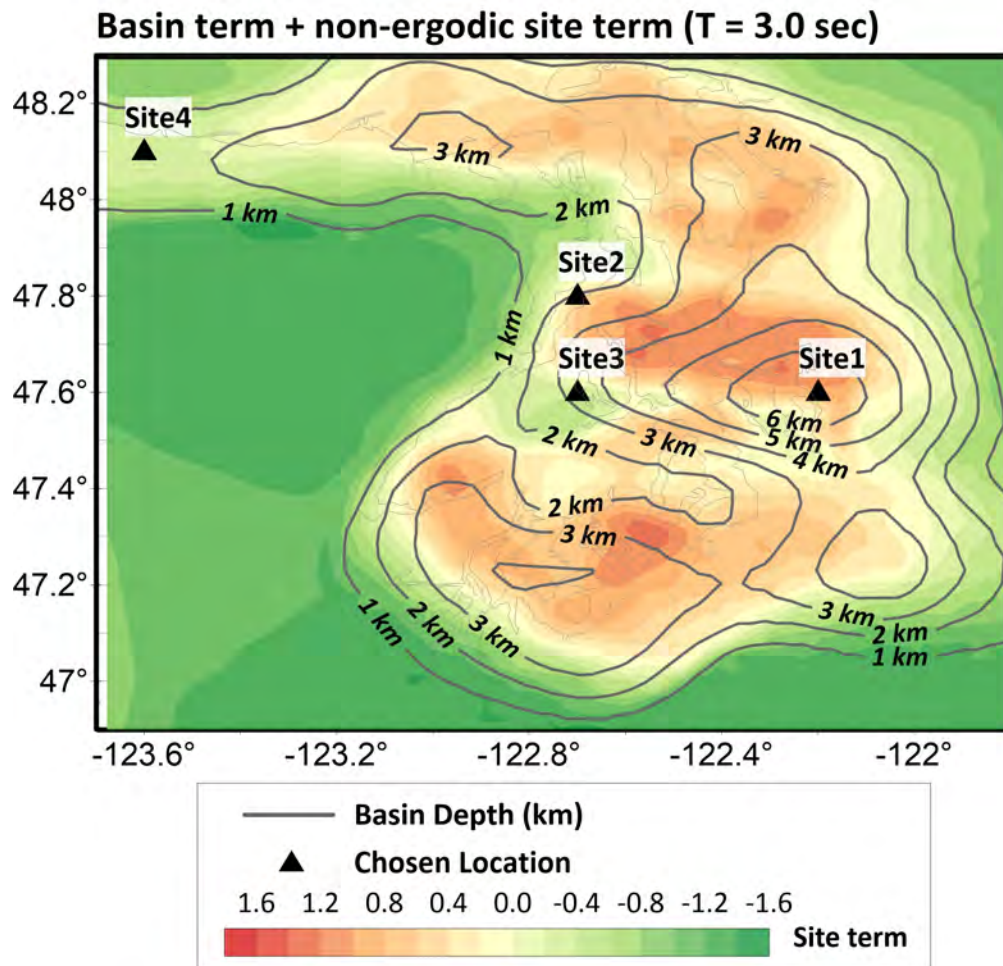


Figure 38: Location of the four sites on the total site term map.

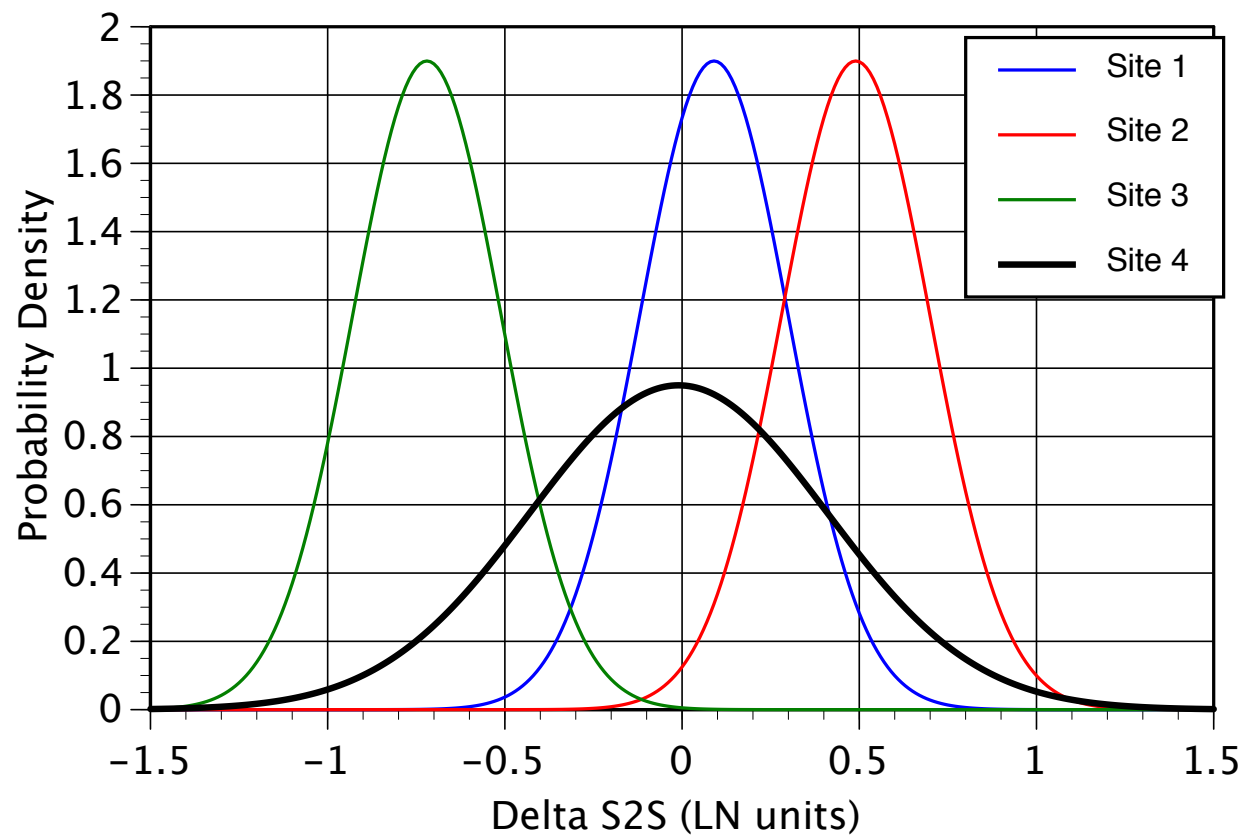


Figure 39: Central estimates and epistemic uncertainty of non-ergodic site terms for $T = 3\text{sec}$ for the four sites.

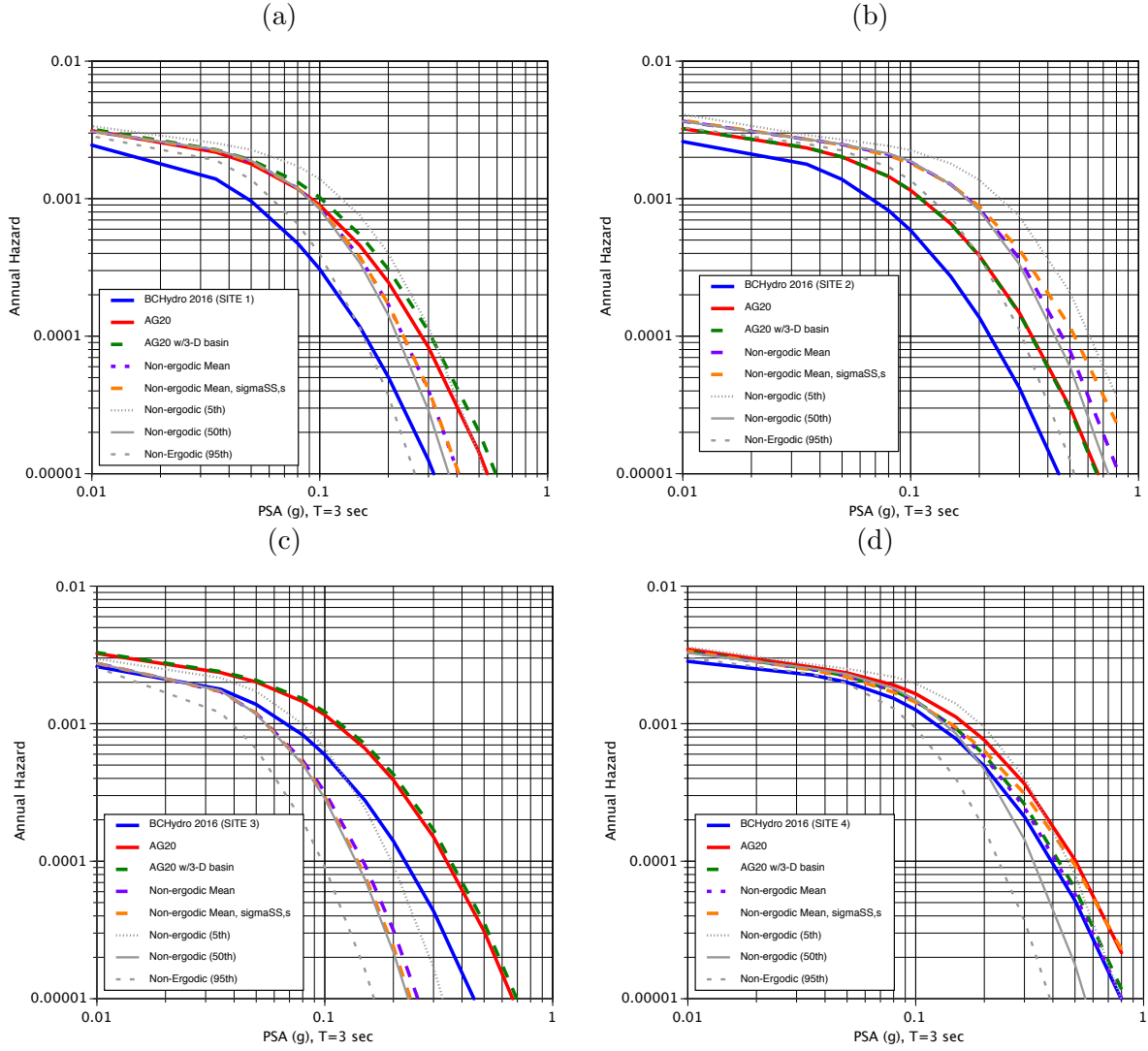


Figure 40: Hazard curves for interface events for different GMM models for $T=3$ sec. (a) Site 1 - very deep basin with negative non-ergodic term. (b) Site 2 - average basin depth with positive non-ergodic term. (c) Site 3 - deep basin with negative non-ergodic site term. (d) Site 4 - shallow basin outside simulation region (zero site term).

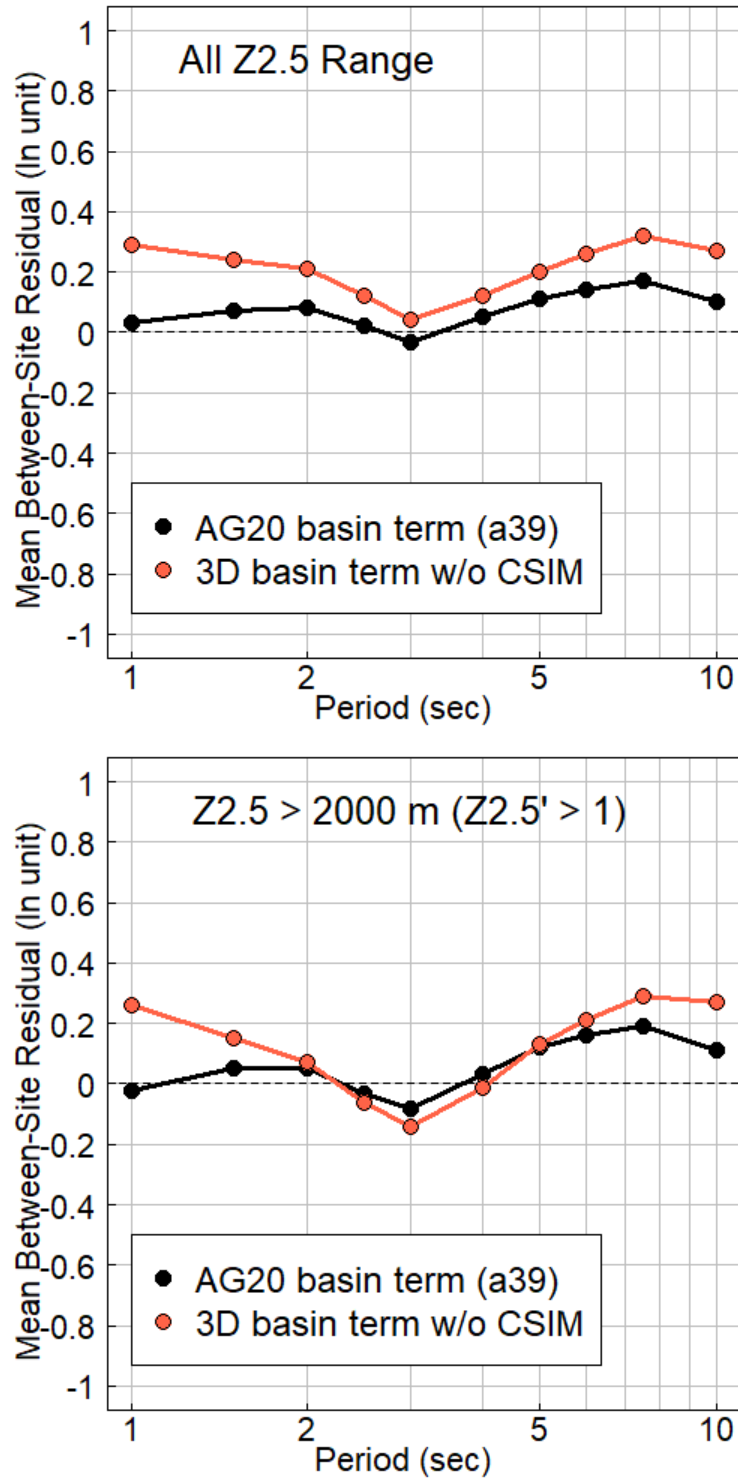


Figure 41: Comparison of mean residuals for intra-slab events using different basin scaling models.

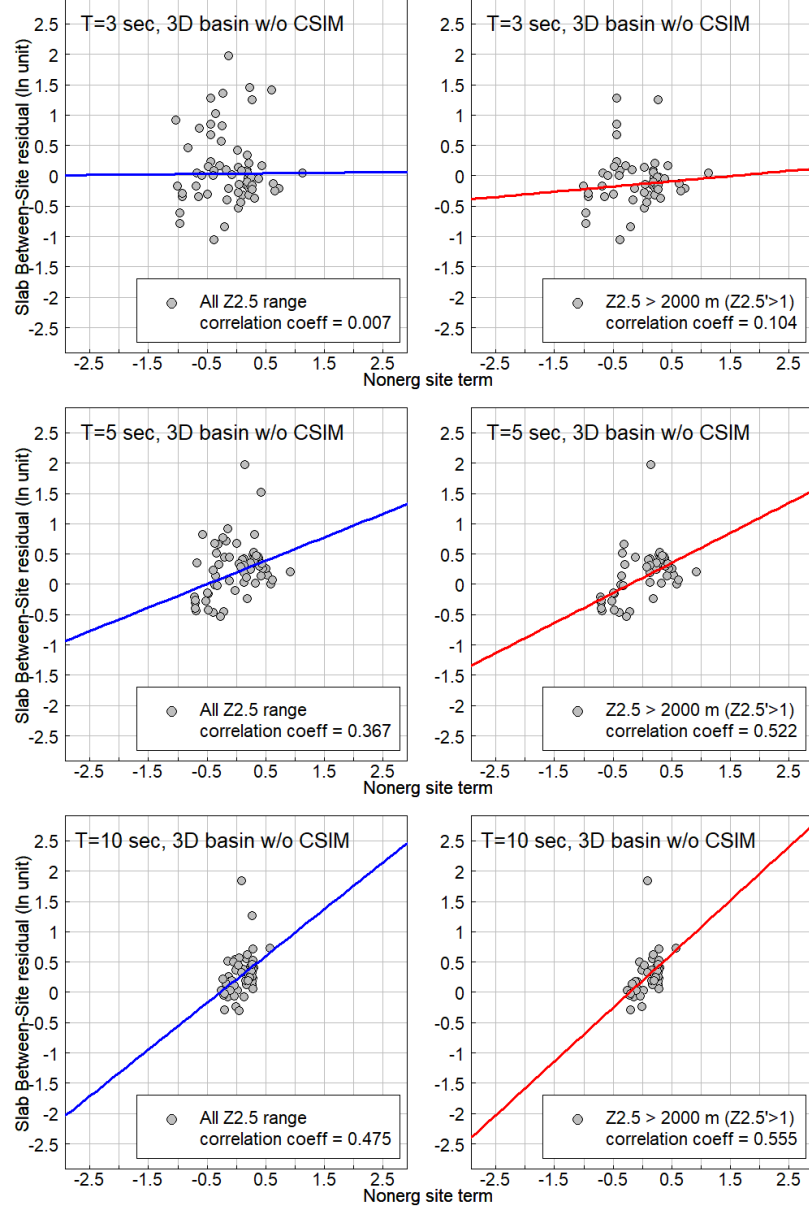


Figure 42: Correlation between site terms for slab events using the modified basin-scaling model with the non-ergodic site terms for interface events from the simulations. (a) T=3 sec. (b) T=5 sec. (c) T=10 sec.

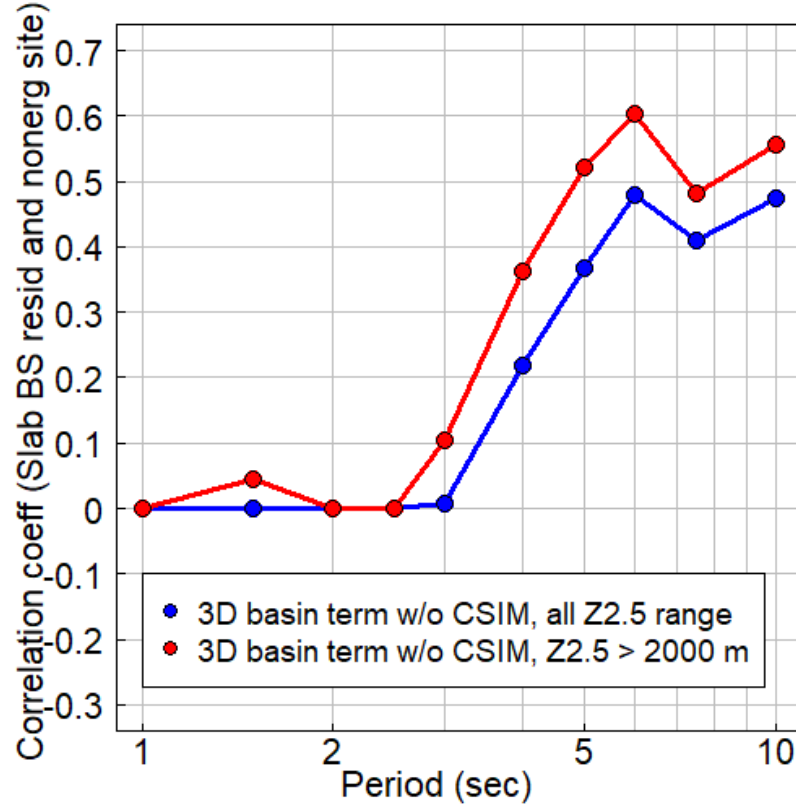


Figure 43: Correlation coefficient of the site terms for slab events using the modified basin-scaling model with the non-ergodic site terms for interface events from the simulations. The blue curve is for all $Z_{2.5}$ values and the red curve is for deeper basin sites with $Z_{2.5} > 2000$ m.